



## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

<b>(51) International Patent Classification 6 :</b> <b>C07H 21/00, 21/02, 21/04</b>	<b>A1</b>	<b>(11) International Publication Number:</b> <b>WO 96/10030</b> <b>(43) International Publication Date:</b> 4 April 1996 (04.04.96)
<b>(21) International Application Number:</b> PCT/US95/13038 <b>(22) International Filing Date:</b> 29 September 1995 (29.09.95) <b>(30) Priority Data:</b> 08/314,877 29 September 1994 (29.09.94) US <b>(60) Parent Application or Grant</b> <b>(63) Related by Continuation</b> US 08/314,877 (CIP) Filed on 29 September 1994 (29.09.94) <b>(71) Applicant (for all designated States except US):</b> ISIS PHARMACEUTICALS, INC. [US/US]; 2280 Faraday Avenue, Carlsbad, CA 92008 (US). <b>(72) Inventors; and</b> <b>(75) Inventors/Applicants (for US only):</b> COOK, Phillip, Dan [US/US]; 641 Avilar Court, San Marcos, CA 92069 (US). TENG, Kelly [US/US]; 4571 Mercurio Street, San Diego, CA 92130 (US). <b>(74) Agents:</b> CALDWELL, John, W. et al.; Woodcock Washburn Kurtz Mackiewicz & Norris, 46th floor, One Liberty Place, Philadelphia, PA 19103 (US).	<b>(81) Designated States:</b> AM, AT, AU, BB, BG, BR, BY, CA, CH, CN, CZ, DE, DK, EE, ES, FI, GB, GE, HU, IS, JP, KE, KG, KP, KR, KZ, LK, LR, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, TJ, TM, TT, UA, UG, US, UZ, VN, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG), ARIPO patent (KE, MW, SD, SZ, UG).  <b>Published</b> <i>With international search report.</i> <i>Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.</i>	

**(54) Title:** OLIGOMERS HAVING HETEROATOM-LINKED CONJUGATE GROUPS**(57) Abstract**

Tetrahydrofuranyl compounds are provided that are functionalized to include pendant conjugate groups, and which are useful in diagnostic assays and as research reagents. Novel intermediates for the synthesis of the compounds are also provided.

BEST AVAILABLE COPY

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	GB	United Kingdom	MR	Mauritania
AU	Australia	GE	Georgia	MW	Malawi
BB	Barbados	GN	Guinea	NE	Niger
BE	Belgium	GR	Greece	NL	Netherlands
BF	Burkina Faso	HU	Hungary	NO	Norway
BG	Bulgaria	IE	Ireland	NZ	New Zealand
BJ	Benin	IT	Italy	PL	Poland
BR	Brazil	JP	Japan	PT	Portugal
BY	Belarus	KE	Kenya	RO	Romania
CA	Canada	KG	Kyrgyzstan	RU	Russian Federation
CF	Central African Republic	KP	Democratic People's Republic of Korea	SD	Sudan
CG	Congo	KR	Republic of Korea	SE	Sweden
CH	Switzerland	KZ	Kazakhstan	SI	Slovenia
CI	Côte d'Ivoire	LI	Liechtenstein	SK	Slovakia
CM	Cameroon	LK	Sri Lanka	SN	Senegal
CN	China	LU	Luxembourg	TD	Chad
CS	Czechoslovakia	LV	Latvia	TG	Togo
CZ	Czech Republic	MC	Monaco	TJ	Tajikistan
DE	Germany	MD	Republic of Moldova	TT	Trinidad and Tobago
DK	Denmark	MG	Madagascar	UA	Ukraine
ES	Spain	ML	Mali	US	United States of America
FI	Finland	MN	Mongolia	UZ	Uzbekistan
FR	France			VN	Viet Nam
GA	Gabon				

## OLIGOMERS HAVING HETEROATOM-LINKED CONJUGATE GROUPS

### CROSS REFERENCE TO RELATED APPLICATIONS

This application is a continuation-in-part of U.S. Serial No. 314,877, filed September 29, 1994, which is a continuation-in-part of U.S. Serial No. 039,846, filed March 30, 1993, which is a continuation-in-part of U.S. Serial No. 903,160, filed June 24, 1992, and PCT Application No. US 92/04294, filed May 21, 1992, which both are continuations-in-part of U.S. Serial No. 703,619, filed May 21, 1991, which is a continuation-in-part of U.S. Serial No. 566,836, filed August 13, 1990, and of U.S. Serial No. 558,663, filed July 27, 1990. This application is also a continuation-in-part of PCT Application No. US 91/05713, filed August 12, 1991, which is a continuation-in-part of U.S. Serial No. 566,836. Each of the foregoing patent applications are assigned to the assignee of this application and are incorporated by reference herein.

### FIELD OF THE INVENTION

This invention is directed to oligomeric compounds that are functionalized to include covalently bound groups. Specifically, the oligomeric compounds include a tetrahydrofuran moiety that is functionalized with a pendant conjugate group.

### BACKGROUND OF THE INVENTION

Oligonucleotides and their analogs have been developed and used in molecular biology in certain procedures as

probes, primers, linkers, adapters, and gene fragments. Modifications to oligonucleotides used in these procedures include labeling with non-isotopic labels, e.g. fluorescein, biotin, digoxigenin, alkaline phosphatase, or other reporter  
5 molecules. Other modifications have been made to the ribose phosphate backbone to increase the nuclease stability of the resulting analog. These modifications include use of methyl and other alkyl phosphonates, phosphorothioates, phosphorodithioate, phosphoamidate and phosphotriester linkages, and  
10 2'-O-methyl ribose sugar units. Further modifications include modification made to modulate uptake and cellular distribution. Phosphorothioate oligonucleotides are presently being used in human clinical trials for various disease states, including use as antiviral agents. In view  
15 of the success of these oligonucleotides for both diagnostic and therapeutic uses, there exists an ongoing demand for improved oligonucleotide analogs.

Oligonucleotides and like molecules can interact with native DNA and RNA in several ways. One of these is duplex  
20 formation between an oligonucleotide and a single stranded nucleic acid. A further method is via triplex formation between an oligonucleotide and double stranded DNA to form a triplex structure.

Naturally occurring or synthetic oligonucleotides, together with hybrid species having both synthetic and  
25 natural components, can collectively be referenced as "oligomeric compounds." Because of their properties, these oligomeric compounds are known to be useful in a number of different areas. They can be used as probes in cloning, blotting procedures, and in applications such as  
30 fluorescence in situ hybridization (FISH). Also, since local triplex formation inhibits gene transcription, such oligomeric compounds can be used to inhibit gene transcription. Labeled oligomers can also be used to  
35 directly map DNA molecules, such as by tagging an oligomer with a fluorescent label and effecting hybridization to complementary sequences in duplex DNA. Oligomers can also

be used as identification tags in combinatorial chemical libraries as is disclosed in patent publication WO 94/08051 and Ohlmeyer et al., *Proc. Natl. Acad. Sci. USA*, 1993, 90, 10922-10926.

5           Considerable research is being directed to the application of oligonucleotides and oligonucleotide analogs that bind complementary DNA and RNA strands for use as diagnostics, research reagents and potential therapeutics.

10           For most uses, it is desirable to append to oligomeric compounds groups that modulate or otherwise influence their activity or their membrane or cellular transport. One method of increasing such transport is by the attachment of a pendant lipophilic group. United States application serial number 117,363, filed Sept. 3, 1993, entitled "Amine-  
15   Derivatized Nucleosides and Oligonucleosides", describes several alkylamino functionalities and their use in the attachment of such pendant groups to oligonucleosides. Additionally, United States application serial number 07/943,516, filed Sept. 11, 1992, entitled "Novel Amines and  
20   Methods of Making and Using the Same" and corresponding to published PCT application WO 94/06815 describe other novel amine-containing compounds, and their incorporation into oligonucleotides for, *inter alia*, the purposes of enhancing cellular uptake, increasing lipophilicity, causing greater  
25   cellular retention and increasing the distribution of the compound within the cell. United States application serial number 08/116,801, filed Sept. 3, 1993, entitled "Thiol-Derivatized Nucleosides and Oligonucleosides", describes nucleosides and oligonucleosides derivatized to include a  
30   thioalkyl functionality, through which pendant groups are attached.

          Although each of the above-noted patent applications describe useful compounds, there remains a need in the art for additional stable compounds that bind complementary DNA  
35   and RNA. There further remains a need in the art for additional methods of attaching pendant groups to oligomeric compounds to further enhance or modulate their binding,

cellular uptake, or other activity.

#### OBJECTS OF THE INVENTION

It is an object of the invention to provide novel nucleosides for use in attaching pendant conjugate groups to oligomeric compounds.

It is a further object of the invention to provide oligomeric compounds that have pendant intercalators, nucleic acid cleaving agents, cell surface phospholipids, diagnostic agents, fluorescent agents and other conjugate groups.

It is yet another object of the invention to provide improvements in research and diagnostic methods and materials for assaying bodily states in animals, especially disease states.

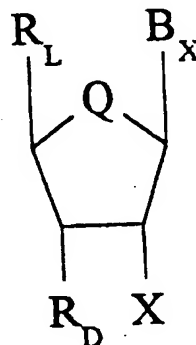
It is a further object of the present invention to provide methods of producing these novel compounds.

These and other objects will become apparent from the following description and accompanying claims.

#### SUMMARY OF THE INVENTION

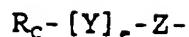
The present invention provides compounds useful in diagnostic assays and as research reagents, as well as methods and intermediates for the preparation thereof.

In one aspect of the invention there are provided compounds of structure:



wherein:

$R_L$  is azido or a group of formula:



in which

Z is O, S or HN;

Y is a bivalent linker;

5 e is 0 or 1;

and

10  $R_c$  is alkyl, alkenyl, alkynyl, O-alkyl, O-alkenyl, O-alkynyl, aryl, alkaryl, phosphinyl, a polyamine, a polyether, polyglycol, a steroid molecule, a reporter molecule, an aromatic or nonaromatic ring system, an aromatic lipophilic molecule, a non-aromatic lipophilic molecule, a reporter enzyme, a peptide, a protein, a water soluble vitamin, a lipid soluble vitamin, a carbohydrate, a terpene molecule, a phospholipid, an intercalator, a cell receptor binding molecule, a crosslinking agent, or a porphyrin;

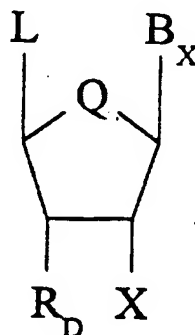
15  $B_x$  is a nucleobase;

X is H, OH, O-alkyl, O-alkoxyalkyl, O-alkylamino or F;

20 Q is O, S,  $CH_2$ , CHF or  $CF_2$ ; and

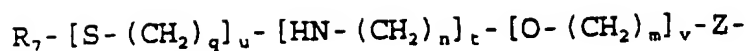
$R_D$  is H, hydroxyl, an activated phosphorous group, a nucleoside, an activated nucleotide, a nucleotide, an oligonucleotide, an oligonucleoside or a protected derivative thereof.

25 Compounds are also provided according to the invention having structure:



wherein:

L is a group of formula:



in which

Z is O, S or HN;

t, u and v are each independently integers from 0 to 200;

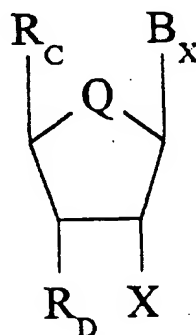
m, n and q are each independently integers from 1 to 4; and

$R_7$  is  $R_c$ , H or a protecting group; and

X,  $B_x$ ,  $R_c$  and  $R_D$  are defined as above.

In some preferred embodiments t and u are 0, or u and v are zero, or t and v are zero. In other preferred embodiments t and u are 0 and m is 2. In other preferred embodiments t and v are 0 and q is 2, or u and v are 0 and n is 2. In even other preferred embodiments u is 0 and each of t and v are 1, and m and n are each 2.

Also provided according to the invention are compounds of structure:



wherein:

$R_c$ ,  $B_x$ , X, Q, and  $R_D$  are defined as above.

In the above structures, in some preferred compounds,  $R_c$  is a polyether or a polyamine. In other preferred embodiments,  $R_c$  is a steroid molecule, preferably cholic acid, deoxycholic acid, dehydrocholic acid, cortisone, digoxigenin, testosterone, cholesterol or 3-trimethylaminomethylhydrazido cortisone.

In some preferred embodiments  $R_c$  is a water soluble vitamin, preferably thiamine, riboflavin, nicotinic acid, pyridoxal phosphate, pyridoxine, pyridoxamine, deoxy-



pyridoxine, pantothenic acid, biotin, folic acid, 5'-deoxyadenosylcobalamin, inositol, choline or ascorbic acid.

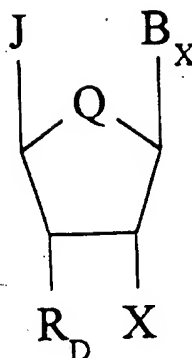
In other preferred embodiments  $R_c$  is a lipid soluble vitamin, preferably a retinal, a retinol, retinoic acid,  $\beta$ -carotene, vitamin D, cholecalciferol, a tocopherol, or a phytol.

In some preferred embodiments  $R_c$  is a protein, preferably a phosphodiesterase, a peroxidase, a phosphatase or a nuclease. In other preferred embodiments  $R_c$  is a reporter molecule, preferably a chromophore, a fluorophore or a radiolabel-containing moiety. Preferred fluorophores include fluorescein, chrysine, anthracene and perylene.

In some preferred embodiments, X is H, OH, F, O-alkyl having from one to six carbons, O-alkylamino having from one to six carbons or O-alkoxyalkyl having from one to six carbons. In certain more preferred embodiments X is H or OH mimicking natural deoxyribo and ribo sugar analogs. In other more preferred embodiments X is F, O-alkyl having from one to six carbons, alkylamino having from one to six carbons or O-alkoxyalkyl having from one to six carbons such that the compounds are homologous to certain nucleotide species that have either greater nuclease resistance or high binding affinity.

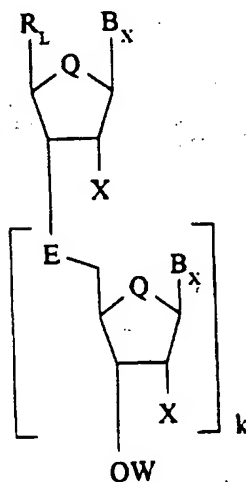
In other preferred embodiments,  $R_p$  is H or OH; or  $R_p$  is an activated phosphorous group, or  $R_p$  is an oligonucleotide.

The present invention also provides novel synthons useful for the preparation of monomeric and polymeric conjugate compounds. In some embodiments these synthons have the structure:



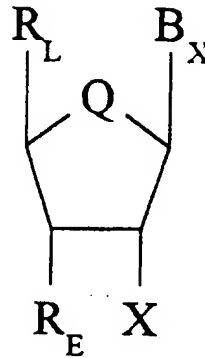
wherein J is a leaving group and Q, R<sub>D</sub>, B<sub>x</sub> and X are as defined above. Preferably, J is OH, SH, NH<sub>2</sub>, trifluoromethylsulfonyl, methylsulfonyl, halogen, O-trichloroacetimidate, acyloxy, dialkyl phosphite, 2,4,6-trichlorophenyl, p-toluenesulfonyl, 4-dimethylaminoazobenzenesulfonyl or 5-dimethylaminonaphthalenesulfonyl.

Also provided are methods for forming a 5'-desmethyl conjugate oligomer having structure:



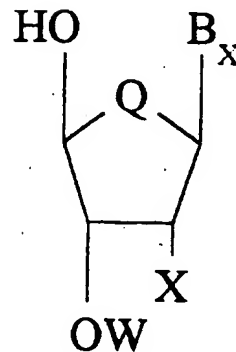
Preferred methods comprise the steps of:

- 10 (a) providing a first synthon having structure:



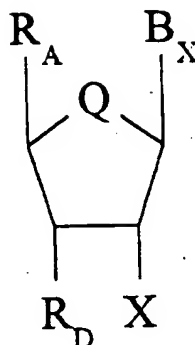
and

(b) contacting the first synthon with a second synthon having structure:

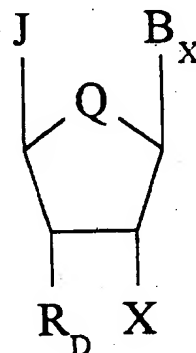


for a time and under reaction conditions sufficient  
 5 to form the conjugated oligomer; wherein  $R_E$  is an activated  
 phosphorous group,  $W$  is  $H$  or a hydroxyl protecting group,  
 $k$  is an integer from 0 to 50,  $E$  is a phosphorous linking  
 species including phosphodiester, phosphotriester,  
 phosphorothioate, phosphorodithioate, alkyl phosphonates  
 10 especially methyl phosphonates and phosphoramidates  
 phosphorous linking groups, and  $Q$ ,  $R_L$ ,  $B_X$  and  $X$  are as  
 defined above.

The present invention also provides methods for  
 forming a 5'-desmethyl conjugated monomer having structure:



In certain embodiments these methods comprise the steps of providing a synthon having structure:



and contacting this synthon with an activated conjugating group. This contacting should be for a time and under reaction conditions sufficient to form the conjugated oligomer. Preferably, J is OH, SH, NH<sub>2</sub>, trifluoromethylsulfonyl, methylsulfonyl, halogen, O-trichloroacetimidate, acyloxy, dialkyl phosphite, 2,4,6-trichlorophenyl, p-toluenesulfonyl, 4-dimethylaminoazobenzenesulfonyl or 5-dimethylaminonaphthalenesulfonyl; more preferably trifluoromethylsulfonyl, methylsulfonyl, halogen or acyloxy.

#### DETAILED DESCRIPTION OF PREFERRED EMBODIMENTS

The present invention provides novel nucleosides for incorporation into nucleic acids. In one aspect, the invention provides nucleosides bearing pendant conjugate groups.

The term "nucleoside" refers to a unit composed of a heterocyclic base and a sugar, generally a pentose sugar. In naturally occurring nucleosides, the heterocyclic base typically is guanine, adenine, cytosine, thymine or uracil, and the sugar is normally deoxyribose, i.e., erythro-pentofuranosyl, or ribose, i.e., ribopentofuranosyl. Synthetic sugars also are known, including arabino, xylo or lyxo pentofuranosyl sugars and hexose sugars. Further synthetic sugar are the 4'-deoxy-4'-thio sugars (both 2'-deoxy and ribo varieties) that have been described by Secrist et al., *J. Med. Chem.* 1991, 34, 2361-2366 and patent application PCT/FR93/00115, respectively, (whereby Q in the above structures is S). Throughout this specification, reference to the sugar portion of a nucleoside or other nucleic acid species shall be understood to refer to naturally occurring deoxyribose and ribose sugars and also to sugar species which can replace the sugar moieties of naturally occurring nucleic acids.

Reference to the nucleobase or heterocyclic base portion of a nucleoside or other nucleic acid species shall be understood to refer to naturally-occurring nucleobases, modified derivatives thereof, or to synthetic nucleobases. Synthetic nucleobases are those which can replace one or more of the naturally occurring base moieties of wild type nucleic acids.

Reference to inter-sugar linkages shall be taken to include moieties that join the sugar portions of nucleosides or nucleotides.

The term "nucleotide" refers to a nucleoside having a phosphate group esterified to one of its 2', 3' or 5' sugar hydroxyl groups. The phosphate group normally is a monophosphate, diphosphate or triphosphate.

The term "oligonucleotide" refers to a plurality of monophosphate nucleotide units that are typically formed from naturally occurring bases and pentofuranosyl sugars joined by native phosphodiester bonds in a specific sequence. A homo-oligonucleotide is formed from nucleotide

units having the same heterocyclic base, e.g. polyadenosine.

The term "oligonucleotide analog" has been used in various published patent application specifications and other literature to refer to molecular species that are similar to oligonucleotides, but that have non-naturally occurring portions. This term has been used to identify oligonucleotide-like molecules that have altered sugar moieties, altered base moieties or altered inter-sugar linkages. Thus, the term oligonucleotide analog has been used to denote structures having altered inter-sugar linkages including phosphorothioate, phosphorodithioate, methyl phosphonate, phosphotriester or phosphoramidate inter-nucleoside linkages in place of phosphodiester inter-nucleoside linkages; purine and pyrimidine heterocyclic bases other than guanine, adenine, cytosine, thymine or uracil and sugars having other than the  $\beta$  pentofuranosyl configuration or sugars having substituent groups at their 2' position or substitutions for one or more of the hydrogen atoms. The term "modified oligonucleotide" also has been used in the literature to denote such structures.

The term "oligonucleotide mimic" has also been used to refer to macromolecular moieties that function similarly to or "mimic" the function of oligonucleotides but have non-naturally occurring inter-sugar linkages. Oligonucleotide mimics thus can have natural or altered or non-naturally occurring sugar moieties and natural or altered or non-naturally occurring base moieties in combination with non-naturally occurring dephospho linkages. Certain dephospho linkages have been reviewed by Uhlmann, E. and Peyman, A., "Oligonucleotide Analogs Containing Dephospho Internucleoside Linkages" in *Methods in Molecular Biology*, Chapter 16, Oligonucleotide Synthetic Protocols, S. Agrawal, Ed., The Humana Press, Inc., Totowa, NJ, 1993.

The term "oligomers" is intended to encompass oligonucleotides, oligonucleotide analogs, oligonucleosides or oligonucleotide-mimicking macromolecules. Thus, "oligomers" refers to nucleosides or nucleoside analogs that

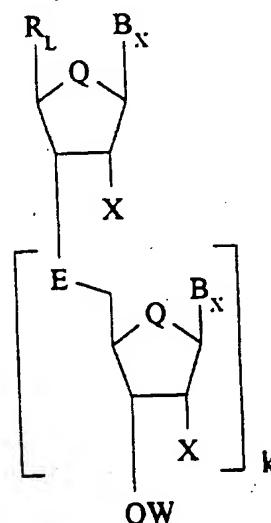
are joined together via either natural phosphodiester bonds or via other linkages.

In certain embodiments the nucleoside compounds of the invention lack the 5'-methylene group present in conventional pentofuranosyl nucleosides. In certain preferred embodiments a heteroatom occupies the position normally occupied by the missing 5'-methylene group. In one sense, these compounds can be considered as 4'-desmethyl pentofuranosyl nucleosides. However, in accordance with IUPAC rules, the lack of a 5'-methylene carbon on the "sugar portion" of the nucleosides of the invention require their designation as tetrahydrofuranyl moieties. Thus, in naming these compounds according to IUPAC rules, as for example in the identification of the structural positions of the compound, established hierarchical or priority nomenclature rules are followed. Accordingly, in embodiments wherein a heteroatom has replaced the 5'-methylene group, that heteroatom and the pendant conjugate group attached thereto takes priority over the heterocyclic base of the nucleoside. In such structures, the tetrahydrofuranyl ring is numbered counterclockwise and the position occupied by the heteroatom (in what would be the 5' position of a conventional nucleoside) is identified as the 2-position. However, in identifying certain of the protons in the NMR spectra, conventional pentofuranosyl nucleoside numbering has been used (except where otherwise noted) for the tetrahydrofuranyl nucleosides.

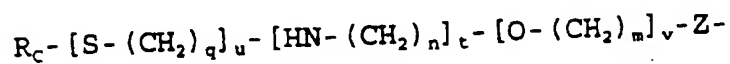
For the purpose of this specification and the claims appended hereto, when an oligomeric structure of the invention is being considered, the ends of this structure are referenced in the same manner as for conventional oligonucleotides. Thus, they are identified either as a 3' end or a 5' end. In other instances where analogy to conventional pentofuranosyl nucleosides is made, strict IUPAC naming rules are deviated from and the numbering system of the conventional pentofuranosyl nucleosides is maintained. In these instances it is convenient to consider

certain tetrahydrofuranyl compounds more as 4'-desmethyl pentofuranosyl compounds and thus attachment is noted as being at the 4' position.

Compounds of the invention include 4'-desmethyl nucleoside monomers modified to bear conjugate groups on their 4'-carbons (i.e., 4'-desmethyl conjugate monomers) or oligomers which contain such modified nucleosides (4'-desmethyl conjugate oligomers). Some oligomeric embodiments have the structure:



wherein a pendant conjugate group  $R_L$  is attached to the 4' position of a nucleoside of the invention. In this structure the "E" group can be any of the commonly use oligonucleotide phosphorous linkage including phosphodiester, phosphotriester, phosphorothioate, phosphorodithioate, alkyl phosphonate and phosphoamidate linkages. The pendant conjugate group can be a group of formula:



wherein Z is O, NH or S, preferably O or NH, t, u and v are each independently integers from 0 to 200; m, n and q are each independently integers from 1 to 4; and  $R_c$  is alkyl, alkenyl, alkynyl, aryl, alkaryl, phosphinyl, O-alkyl, O-



alkenyl, O-alkynyl, a polyglycol, a polyamine, a polyether, a steroid molecule, an aromatic or nonaromatic ring system, a reporter molecule, an aromatic lipophilic molecule, a non-aromatic lipophilic molecule, a reporter enzyme, a peptide, a protein, a water soluble vitamin, a lipid soluble vitamin, a carbohydrate, a terpene molecule, a phospholipid, an intercalator, a cell receptor binding molecule, a crosslinking agent, an RNA cleaving complex, a metal chelator, an alkylator or a porphyrin.

10 In oligomers of the invention at least one 4'-desmethyl nucleoside unit incorporated therein has been derivatized to bear a pendant conjugate group as described above. When so derivatized, the oligomeric compound is useful, for example, in a diagnostic or other preparation that includes a nucleic acid binding agent to assist in identification of the oligomeric compound, to aid in transfer of the compound across cellular membranes, or to impart other properties. Such a diagnostic or nucleic acid binding agent is formed from an oligomeric compound of the invention wherein the oligomeric compound includes monomeric units bearing natural or non-natural occurring bases in a sequence that is complementary to and will specifically hybridize with a region of an RNA or DNA of interest.

20 For the purpose of identification, a functionalized oligomeric compound according to the invention can be characterized as a substituent-bearing (e.g., steroid-bearing) oligomeric compound. Such oligomeric compounds will have at least one pendant group attached thereto to modulate their activity.

30 The compounds of the inventions have various uses including being useful as research reagents and in diagnostic assays. In one particular diagnostic assay, compounds of the inventions are used to isolate the effects of one cellular adhesion molecule from those of a further cellular adhesion molecule in a protein expression assay. In this assay, the effect of induced expression of intercellular adhesion molecule-1 (ICAM-1) were diminished

while those of VCAM-1 were maintained. This allows for analysis of the VCAM-1 protein expression without concurrent expression and interference of ICAM-1 protein.

5 In general, the oligomeric compounds bearing pendant conjugate groups of the present invention can be used to bind to various other target molecules. Target molecules of the present invention can include any of a variety of biologically significant molecules. Other such target molecules can be nucleic acids, carbohydrates, glycoproteins  
10 or other proteins.

In binding to nucleic acids, the functionalized oligomeric compounds of the invention bind with complementary strands of RNA or DNA. After binding, the oligomeric compound and the RNA or DNA strand can be  
15 considered to be complementary strands which are "duplexed" in a manner analogous to native, double-stranded DNA. In such complementary strands, the individual strands are positioned in such a manner with respect to one another so as to allow Watson-Crick type, Hoogsteen type or another type  
20 hybridization of the heterocyclic bases of one strand to the heterocyclic bases of the opposing strand.

Binding to nucleic acids can be practiced against nucleic acids from a variety of sources including organisms ranging from unicellular prokaryotes and eukaryotes to  
25 multicellular eukaryotes. The nucleic acid from any organism that utilizes DNA-RNA transcription or RNA-protein translation as a fundamental part of its hereditary, metabolic or cellular control is susceptible to such binding. Seemingly diverse organisms such as bacteria,  
30 yeast, virus, protozoa, algae, all plant and all higher animal forms, including warm-blooded animals, are sources of such nucleic acid. Further, since each of the cells of multicellular eukaryotes includes both DNA-RNA transcription and RNA-protein translation as an integral part of their  
35 cellular activity, nucleic acid binding for various purposes, e.g. diagnostics, can also be practiced on such cellular populations. Furthermore, many of the organelles,

e.g., mitochondria and chloroplasts, of eukaryotic cells include transcription and translation mechanisms. Thus, single cells, cellular populations or organelles can also be included within the definition of organisms that are

5 capable of being treated with nucleic acid binders. In some preferred embodiments of the present invention, the target molecule is a protein such as an immunoglobulin, receptor, receptor binding ligand, antigen or enzyme and more specifically can be a phospholipase, tumor necrosis factor,

10 endotoxin, interleukin, plasminogen activator, protein kinase, cell adhesion molecule, lipoxxygenase, hydrolase or transacylase. In other embodiments of the invention the target molecules can be important regions of the human immunodeficiency virus, *Candida*, herpes viruses,

15 papillomaviruses, cytomegalovirus, rhinoviruses, hepatitis viruses, or influenza viruses. In further embodiments, the target molecules can be regions of an oncogene. In still further embodiments, the target molecule is ras 47-mer stem loop RNA, the TAR element of human immunodeficiency virus

20 or the gag-pol stem loop of human immunodeficiency virus (HIV). Still other targets can induce cellular activity. For example, a target can induce interferon.

Binding also can be practiced against transcription factors. In binding to transcription factors or other

25 target molecules, the transcription factor or other target molecule need not be purified. It can be present, for example, in a whole cell, in a humoral fluid, in a crude cell lysate, in serum or in other humoral or cellular extracts. Of course, purified transcription factor or a

30 purified form of another target molecule is also useful in some aspects of the invention.

In still other embodiments of the present invention, synthetically prepared transcription factors or other target molecules can be useful. A transcription factor or other

35 target molecule also can be modified, such as by biotinylation or radiolabeling. For example, a synthetically prepared transcription factor can incorporate one or more

biotin molecules during synthesis or can be modified post-synthesis.

Transcription factors, as the term is used herein, are DNA- or RNA-binding proteins that regulate the expression of genes. HIV tat and c-rel are examples of transcription factors which regulate the expression of genes. Also encompassed by the term are DNA and RNA binding proteins which are not strictly considered transcription factors, but which are known to be involved in cell proliferation. These transcription factors include c-myc, fos, and jun. Methods of the present invention are particularly suitable for use with transcription factors as target molecules since transcription factors generally occur in very small cellular quantities.

In still other embodiments of the invention, nucleic acid binding can be used to treat objects (glasswares, petri dishes, instruments or the like) to sterilize, sanitize or disinfect such objects that may harbor an organism, e.g., bacterial, protozoan, viral, fungal or other infectious or non-infectious agent on the object. In binding to the nucleic acid of the agent, the compounds of the invention will kill, abrogate, inhibit, curtail or otherwise control, eradicate or render harmless the cellular growth or gene expression of such organisms on the object. Such nucleic acid binding can also be used to render such organisms incapable or otherwise unable to reproduce themselves.

In still further embodiments of the inventions, the nucleic acid binding properties of the compounds of the invention can be used to form duplex structures with an unknown oligomer species for identification of that species by gel analysis including slab and capillary gel electrophoresis. In still further embodiments of the inventions, the compounds of the invention can be used as tags in identification of other compounds in combinatorial libraries.

Preferred nucleobase units for oligomeric compounds of the invention include naturally occurring or synthetic

purine or pyrimidine heterocyclic bases, including but not limited to adenine, guanine, cytosine, thymine, uracil, 5-methylcytosine, hypoxanthine or 2-aminoadenine. Other such heterocyclic bases include 2-methylpurine, 2,6-diaminopurine, 6-mercaptapurine, 2,6-dimercaptapurine, 2-amino-6-mercaptapurine, 5-methylcytosine, 4-amino-2-mercaptopyrimidine, 2,4-dimercaptopyrimidine and 5-fluorocytosine. Representative heterocyclic bases are disclosed in U.S. Patent No. 3,687,808 (Merigan, et al.), which is incorporated herein by reference.

In preferred embodiments, the oligomeric compounds of the invention include from about 2 to about 50 nucleoside subunits (i.e.,  $k$  = about 1 to about 49).

As used in this specification, alkyl groups of the invention include but are not limited to  $C_1$ - $C_{12}$  straight and branched chained alkyls such as methyl, ethyl, propyl, butyl, pentyl, hexyl, heptyl, octyl, nonyl, decyl, undecyl, dodecyl, isopropyl, 2-butyl, isobutyl, 2-methylbutyl, isopentyl, 2-methylpentyl, 3-methylpentyl, 2-ethylhexyl and 2-propylpentyl. Alkenyl groups include but are not limited to unsaturated moieties derived from the above alkyl groups including but not limited to vinyl, allyl and crotyl. Alkynyl groups include unsaturated moieties having at least one triple bond that are derived from the above alkyl groups including but are not limited to ethynyl and propargyl. Alkanoyl groups according to the invention are alkyl, alkenyl or alkynyl groups attached through a carbonyl group.

The term aryl is intended to denote monocyclic and polycyclic aromatic groups including, for example, phenyl, naphthyl, xylyl, pyrrole, and furyl groups. Although aryl groups (e.g., imidazo groups) can include as few as 3 carbon atoms, preferred aryl groups have 6 to about 14 carbon atoms, more preferably 6 to about 10 carbon atoms. Aralkyl and alkaryl groups according to the invention each include alkyl and aryl portions. Aralkyl groups are attached through their alkyl portions, and alkaryl groups are attached through their aryl portions. Benzyl groups provide

one example of an aralkyl group, and p-tolyl provides an example of an alkaryl group.

Alkylamino and aminoalkyl groups according to the invention each include amino and alkyl portions. Alkylamino groups are attached through their amino portions, and aminoalkyl groups are attached through their alkyl portions. Methylamino groups provide one example of an alkylamino group, a  $\beta$ -aminobutyl group is one example of an aminoalkyl group.

For the purposes of this invention a ring system is defined to include two or more single rings that join together to form an extended or condensed ring. Such ring systems include extended aromatic systems, alicyclic systems, araalicyclic systems, and bicyclic systems. Examples include aromatic, alicyclic and mixed aromatic-alicyclic (araalicyclic) multiple ring systems, bicyclic systems, non-aromatic multiple ring systems such as adamantane, decalin, and norbornane. Ring systems also include, but are not limited to, naphthalene, tetrahydronaphthalene (tetralin), anthracene, phenanthrene, fluorene, pyrene, coronene, azulene, fluorene, benzonaphthene, benzo[8]annulene, pentalene, heptalene, octalene, indene, isoindene biphenyl, biphenylene and triphenylene condensed rings.

The terms alkyl, alkaryl, aralkyl and aryl are intended to denote both substituted (e.g., halogenated and hydroxylated) and unsubstituted moieties.

Halogens include fluorine, chlorine and bromine. Suitable heterocyclic groups include but are not limited to imidazole, tetrazole, triazole, pyrrolidine, piperidine, piperazine and morpholine. Heterocycloalkyl groups are cyclic alkyl groups containing a heretoatom. Heterocycloalkaryl groups are aryl heterocycles bearing at least one alkyl substituent.

Amines include amines of all of the above alkyl, alkenyl and aryl groups including primary and secondary amines and "masked amines" such as phthalimide. Amines are

also meant to include polyalkylamino compounds and aminoalkylamines such as aminopropylamine and further heterocycloalkylamines such as imidazol-1, 2 or 4-yl-propylamine.

5           Substituent groups for the above as well as for other moieties listed below include but are not limited to other alkyl, haloalkyl, alkenyl, alkynyl, alkoxy, thioalkoxy, haloalkoxy and aryl groups as well as halogen, hydroxyl, amino, azido, carboxy, cyano, nitro, mercapto, sulfides, 10   sulfones, sulfoxides, keto, carboxy, nitrates, nitrites, nitroso, nitrile, trifluoromethyl, O-alkyl, S-alkyl, NH-alkyl, amino, silyl, amides, ester, ethers, carbonates, carbamates, ureas, imidazoles, intercalators, conjugates, polyamines, polyamides, polyethylene glycols, polyethers, 15   groups that enhance the pharmacodynamic properties of oligonucleotides, and groups that enhance the pharmacokinetic properties of oligonucleotides. Other suitable substituent groups also include rhodamines, coumarins, acridones, pyrenes, stilbenes, oxazolopyridocarbazoles, 20   anthraquinones, phenanthridines, phenazines, azidobenzenes, psoralens, porphyrins and cholesterol. One particularly preferred group is  $\text{CF}_3$ .

As used in the present invention, the term polyether means a linear or branched alkyl chain periodically 25   interrupted by oxygen atoms. One preferred example of a polyethers is polyethylene glycol. The term polyglycol, as used herein, includes compounds such as glycerol, and derivative groups such as 2,3-dialkoxypropyl groups. The term polyamine as used herein includes the nitrogen analogs 30   of such structures, and the term polythioether includes the sulfur analogs of such structures.

Carbohydrates according to the invention are inclusive of pentose, hexose and higher sugars, and the polymeric species thereof. Representative carbohydrates include 35   glucose and galactose and their derivatives including, as for example, glycals, glycal epoxides and glycosides.

Terpenes are known in the art as oligomers of

isoprene, particularly the dipentenenes, pinenes, and myrcenes. Included within the definition of terpene molecules are terpene derivatives such as camphor and menthol.

5 The term phospholipid as used herein includes those compounds which upon hydrolysis yield phosphoric acid, an alcohol and one or more fatty acids. Representative examples of phospholipids include lecithin, cephalin and sphingomyelin.

10 As used in the present invention, groups that enhance the pharmacodynamic properties include groups that improve oligonucleotide uptake, enhance oligonucleotide resistance to degradation, and/or strengthen sequence-specific hybridization with RNA. Groups that enhance the pharmaco-  
15 kinetic properties, in the context of this invention, include groups that improve oligonucleotide uptake, distribution, metabolism or excretion.

For the purposes of this invention, the terms "reporter molecule" and "reporter enzyme" are inclusive of  
20 those molecules or enzymes that have physical or chemical properties that allow them to be identified in gels, fluids, whole cellular systems, broken cellular systems and the like utilizing physical properties such as spectroscopy, radioactivity, colorimetric assays, fluorescence, and  
25 specific binding. Particularly useful as reporter molecules are fluorophores, chromaphores and radiolabel-containing moieties. Fluorophores are molecules detectable by fluorescence spectroscopy. Examples of preferred fluorophores are fluorescein and rhodamine dyes and  
30 acridines. There are numerous commercial available fluorophores including "Texas Red" and other like fluoresceins and rhodamines available for Molecular Probes, Eugene, OR. Chromaphores are molecules capable of detection by visible or ultraviolet (UV-VIS) absorbance spectroscopy.  
35 Examples of chromaphores are polynuclear aromatics such as anthracene, perylene, pyrene, rhodamine and chrysene. Radiolabel-containing moieties, as used herein, are



molecules incorporating at least one radioactive atom, such as  $^3\text{H}$  or  $^{14}\text{C}$ , enabling detection thereby. Reporter enzymes may be detected directly or via their enzymatic products by any of the methods mentioned above. Particularly useful as  
5 reporter enzymes are alkaline phosphatase and horseradish peroxidase.

Steroid molecules according to the invention include those chemical compounds that contain a perhydro-1,2-cyclopentanophenanthrene ring system. Particularly useful  
10 as steroid molecules are the bile acids including cholic acid, deoxycholic acid and dehydrocholic acid; steroids including cortisone, digoxigenin, testosterone and cholesterol and cationic steroids such as cortisone having a trimethylaminomethyl hydrazide group attached via a double  
15 bond at the 3-position of the cortisone ring (3-trimethylaminomethylhydrazido cortisone).

Proteins and peptides are utilized in their usual sense as polymers of amino acids. Normally peptides comprise such polymers that contain a smaller number of  
20 amino acids per unit molecule than do the proteins. Particularly useful as peptides and proteins are sequence-specific peptides and proteins including phosphodiesterases, peroxidases, phosphatases and nucleases. Such peptides and proteins include, but are not limited to, SV40 peptide,  
25 RNase A, RNase H and Staphylococcal nuclease.

Lipophilic molecules include naturally-occurring and synthetic aromatic and non-aromatic moieties such as fatty acids, esters and alcohols, other lipid molecules, cage structures such as adamantane and buckminsterfullerenes, and  
30 aromatic hydrocarbons such as benzene, perylene, phenanthrene, anthracene, naphthalene, pyrene, chrysene, and naphthacene. Particularly useful as lipophilic molecules are alicyclic hydrocarbons, saturated and unsaturated fatty acids, waxes, terpenes and polyalicyclic hydrocarbons  
35 including adamantane and buckminsterfullerenes.

Alkylators according to the invention are moieties that can effect attachment of electrophilic groups to

targeted molecular structures. Representative alkylators are disclosed by Meyer, et al., *J. Am. Chem. Soc.* 1989, 111, 8517.

Intercalators are polycyclic aromatic moieties that can insert between adjacent base pairs without affecting normal Watson-Crick base pairing, and include hybrid intercalator/ligands such as the photonuclease/intercalator ligand 6-[[[9-[[6-(4-nitrobenzamido)hexyl]amino]acridin-4-yl]carbonyl]amino]hexanoylpentafluorophenyl ester. This compound has two noteworthy features: an acridine moiety that is an intercalator and a p-nitrobenzamido group that is a photonuclease. Other representative intercalators are disclosed by Manoharan, M., *Antisense Research and Applications*, Crooke and Lebleu, eds., CRC Press, Boca Raton, 1993.

Cell receptor binding molecules according to the invention are vitamins and carbohydrate moieties for which specific receptors exist within a cell. Representative cell receptor binding molecules are disclosed by Application Serial No. PCT/US92/09196, filed October 23, 1992, the contents of which are incorporated herein by reference.

Crosslinking agents are moieties that can effect intrastrand or interstrand covalent binding of RNA and/or DNA, and include photo-crosslinking agents. Examples of photocrosslinking agents include aryl azides such as N-hydroxysuccinimidyl-4-azidobenzoate (HSAB) and N-succinimidyl-6(-4'-azido-2'-nitrophenylamino)hexanoate (SANPAH). Aryl azides conjugated to oligomers will effect crosslinking with nucleic acids and proteins upon irradiation. Other representative crosslinking agents are disclosed in International Patent Application Serial No. PCT/US93/02059, filed March 5, 1993, which is incorporated herein by reference.

Useful crown amines are disclosed by Studer, et al., *Helv. Chim. Acta* 1986, 69, 2081 and Smith-Jones, et al., *Bioconjugate Chem.* 1991, 2, 415.

Vitamins according to the invention generally can be

classified as water soluble or lipid soluble. Water soluble vitamins include thiamine, riboflavin, nicotinic acid or niacin, the vitamin B<sub>6</sub> pyridoxal group, pantothenic acid, biotin, folic acid, the B<sub>12</sub> cobalamin coenzymes, inositol, 5 choline and ascorbic acid. Lipid soluble vitamins include the vitamin A family, vitamin D, the vitamin E tocopherol family and vitamin K (and phytols). The vitamin A family, including retinoic acid and retinol, are absorbed and transported to target tissues through their interaction with 10 specific proteins such as cytosol retinol-binding protein type II (CRBP-II), Retinol-binding protein (RBP), and cellular retinol-binding protein (CRBP). These proteins, which have been found in various parts of the human body, have molecular weights of approximately 15 kD. They have 15 specific interactions with compounds of the vitamin A family, especially, retinoic acid and retinol.

The vitamin A family of compounds can be attached to oligomers of the invention via acid or alcohol 20 functionalities found in the various family members. For example, conjugation of an N-hydroxysuccinimide ester of an acid moiety of retinoic acid to an amine function of a pendant group of the oligomer results in linkage of the vitamin A compound to the oligomer, via an amide bond. In 25 similar fashion, standard esterification chemistries may be used to attach the acid moiety of the retinoic acid group to a 4'-oxygen of a compound of the invention, or to a hydroxyl function of a pendant group thereof.

$\alpha$ -Tocopherol (vitamin E) and other tocopherols (beta through zeta) can be similarly conjugated to oligomers also 30 to enhance uptake due to their lipophilic character. The lipophilic vitamin, vitamin D, and its ergosterol precursors can be conjugated to oligomers through their hydroxyl groups by first activating the hydroxyl groups by forming, for example, hemisuccinate esters. Conjugation then is 35 effected, as for instance, to an aminolinker pendant from the oligomer, or through other suitable functional groups described herein. Other vitamins that can be conjugated to

oligomers through hydroxyl groups on the vitamins include thiamine, riboflavin, pyridoxine, pyridoxamine, pyridoxal, deoxypyridoxine. Lipid soluble vitamin K's and related quinone-containing compounds can be conjugated via carbonyl groups on the quinone ring. The phytol moiety of vitamin K may also serve to enhance binding of the oligomers to cells.

Pyridoxal phosphate is the coenzyme form of Vitamin B<sub>6</sub>. Vitamin B<sub>6</sub> is also known as pyridoxine. Pyridoxal has specific B<sub>6</sub>-binding proteins. The role of these proteins in pyridoxal transport has been studied by Zhang and McCormick, *Proc. Natl. Acad. Sci. USA*, 1991 88, 10407. Zhang and McCormick showed that a series of N-(4'-pyridoxyl)amines, in which several synthetic amines were conjugated at the 4'-position of pyridoxal, were able to enter cells by a process facilitated by the B<sub>6</sub> transporter. Zhang and McCormick also demonstrated the release of these synthetic amines within the cell. Other pyridoxal family members include pyridoxine, pyridoxamine, pyridoxal phosphate, and pyridoxic acid. Pyridoxic acid, niacin, pantothenic acid, biotin, folic acid and ascorbic acid can be conjugated to oligomers using N-hydroxysuccinimide esters as described above for retinoic acid.

Other groups for modifying properties of oligomers include RNA cleaving complexes, pyrenes, metal chelators, porphyrins, alkylators, hybrid intercalator/ligands and photo-crosslinking agents. RNA cleavers include o-phenanthroline/Cu complexes and Ru(bipyridine)<sub>2</sub><sup>2+</sup> complexes. The Ru(bipyridine)<sub>2</sub><sup>2+</sup> complexes interact with nucleic acids and cleave nucleic acids photochemically. Metal chelators include EDTA, DTPA, and o-phenanthroline. Alkylators include compounds such as iodoacetamide. Porphyrins include porphine, its substituted forms, and metal complexes. Pyrenes include pyrene and other pyrene-based carboxylic acids that can be conjugated using protocols similar to those specified above.

Numerous suitable protecting groups are known in the

art for protecting the several functional groups of the compounds of the invention during synthesis. Protecting groups are known per se as chemical functional groups that can be selectively appended to and removed from functionalities, such as hydroxyl groups and carboxyl groups. These groups are present in a chemical compound to render such functionality inert to chemical reaction conditions to which the compound is exposed. Protecting groups useful in the context of the present invention include but are not limited to hydroxyl protecting groups such as t-butyldiphenylsilyl, t-butyldimethylsilyl, and dimethoxytrityl groups, thiol protecting groups such as S-trityl, S-p-methoxybenzylthioether, S-p-nitrobenzylthioether, and S-t-butylthioether. (see, e.g., Veber and Hirschmann, et al., *J. Org. Chem.* 1977, 42, 3286 and Atherton, et al., *The Peptides*, Gross and Meienhofer, Eds, Academic Press; New York, 1983; Vol. 9 pp. 1- 38). Other representative protecting groups suitable for practice in the invention may be found in Greene, T.W. and Wuts, P.G.M., "Protective Groups in Organic Synthesis" 2d. Ed., Wiley & Sons, 1991.

In one aspect, the present invention is directed to nucleoside monomers that bear at least one conjugate group at the 4'-position. In oligomeric embodiments, the conjugate-bearing nucleoside is the 5'-terminal residue oligomer. In some preferred embodiments the conjugate group is a polyamine, and in other preferred embodiments the conjugate group is a polyether. In further preferred embodiments the conjugate group is a polythioether.

The conjugate groups of the preceding paragraph can further serve as bivalent linkers for other conjugate groups. Thus for instances, a hydroxyl, amino or thiol terminated polyether, polyamine or polythioether is used a linker and the terminal group is then reacted with a suitable functional group on a conjugate molecules to link the conjugate molecule to the polyether, polyamine or polythioether. Suitable functional groups for reacting with

hydroxyl, amine or thiol groups include a gamut of groups known in the organic chemical arts including, but not limited to, acid chlorides, anhydrides, cyclic anhydrides, alkyl halides, organometallics, chloroformates, isocyanates and the like. In a like manner other terminal groups can be utilized on the linker including, but not limited to, hydrazines, aldehydes, acids, hydroxylamines, semicarbazides, thiosemicarbazides, hydrazones, hydrazides and the like. Protected forms of such compounds can also be used, as for example, protected aldehydes including, C-formyl, o-methylaminobenzenethio, aryl substituted imidazolidino moieties and the like. In a like manner, protected thiol include trityl thiol that can be deblocked allowing a disulfide bonds to be formed between the linker and the a sulfur bearing conjugate group to join the conjugate group to the linker. Various nitrogen linkages can be used to join the conjugate groups including hydrazones, oximes, hydrazide-hydrazone, semicarbazone and semithiocarbazones.

The phrase polyamine species, including polyamines, as used herein refers to species that have a plurality of nitrogen atoms thereon. Polyamines include primary amines, hydrazines, semicarbazines, thiosemicarbazines and similar nitrogenous species. Such species can be symmetrical species such as polyamine containing polymers or they can be unsymmetrical wherein the amine functionalities of the polyamine are separated in space by different moieties. In addition to carbon atoms other atomic species such as nitrogen and sulfur may also be incorporated into the polyamine species. In some preferred embodiments of the invention, at least one nitrogen atom of the polyamine has a free electron pair.

Preferred as polyamine species are species that range in length from about 3 to about 20 units. More preferably species having at least one nitrogen atom and are of the general formula  $H_2N[(CH_2)_n-(NH)_t]_x$  wherein n is an integer between 2 and 8 and t is an integer between 1 and 10. These

species can be linear or cyclic. Cyclic amines would include crown amines and mixed crown amines/crown ethers.

Other suitable nitrogen-containing compounds suitable for the formation of polyamine species include

5 straight chain alkylamine,  $C_1-C_{20}$  straight chain substituted alkylamine,  $C_2-C_{50}$  branched chain alkylamine,  $C_2-C_{50}$  branched chain substituted alkylamine,  $C_3-C_{50}$  cyclic alkylamine,  $C_3-C_{50}$  cyclic substituted alkylamine,  $C_2-C_{20}$  straight chain alkenylamine,  $C_2-C_{20}$  straight chain substituted alkenylamine,

10  $C_3-C_{50}$  branched chain alkenylamine,  $C_3-C_{50}$  branched chain substituted alkenylamine,  $C_3-C_{50}$  cyclic alkenylamine,  $C_3-C_{50}$  cyclic substituted alkenylamine,  $C_2-C_{20}$  straight chain alkynylamine,  $C_2-C_{20}$  straight chain substituted alkynylamine,  $C_3-C_{50}$  branched chain alkynylamine,  $C_3-C_{50}$  branched chain

15 substituted alkynylamine,  $C_3-C_{50}$  cyclic alkynylamine,  $C_3-C_{50}$  cyclic substituted alkynylamine,  $C_1-C_{20}$  straight chain alkylhydrazine,  $C_1-C_{50}$  straight chain substituted alkylhydrazine,  $C_2-C_{50}$  branched chain alkylhydrazine,  $C_2-C_{50}$  branched chain substituted alkylhydrazine,  $C_3-C_{50}$  cyclic

20 hydrazoalkane,  $C_3-C_{50}$  cyclic substituted hydrazoalkane,  $C_2-C_{20}$  straight chain alkenylhydrazine,  $C_2-C_{20}$  straight chain substituted alkenylhydrazine,  $C_3-C_{50}$  branched chain alkenylhydrazine,  $C_3-C_{50}$  branched chain substituted alkenylhydrazine,  $C_3-C_{50}$  cyclic hydrazoalkene,  $C_3-C_{50}$  cyclic

25 substituted hydrazoalkene,  $C_2-C_{20}$  straight chain alkynylhydrazine,  $C_2-C_{20}$  straight chain substituted alkynylhydrazine,  $C_3-C_{50}$  branched chain alkynylhydrazine,  $C_3-C_{50}$  branched chain substituted alkynylhydrazine,  $C_3-C_{50}$  cyclic hydrazoalkyne,  $C_3-C_{50}$  cyclic substituted hydrazoalkyne,  $C_1-C_{20}$

30 straight chain alkylhydroxyamine,  $C_1-C_{20}$  straight chain substituted alkylhydroxyamine,  $C_2-C_{50}$  branched chain alkylhydroxyamine,  $C_2-C_{50}$  branched chain substituted alkylhydroxyamine,  $C_3-C_{50}$  cyclic oxyalkylamine,  $C_3-C_{50}$  cyclic substituted oxyalkylamine,  $C_2-C_{20}$  straight chain

35 alkenylhydroxyamine,  $C_2-C_{20}$  straight chain substituted alkenylhydroxyamine,  $C_3-C_{50}$  branched chain alkenylhydroxyamine,  $C_3-C_{50}$  branched chain substituted

- alkenylhydroxyamine, C<sub>3</sub>-C<sub>50</sub> cyclic oxyalkenylamine, C<sub>3</sub>-C<sub>50</sub>  
cyclic substituted oxyalkenylamine, C<sub>2</sub>-C<sub>20</sub> straight chain  
alkynylhydroxyamine, C<sub>2</sub>-C<sub>20</sub> straight chain substituted  
alkynylhydroxyamine, C<sub>3</sub>-C<sub>50</sub> branched chain  
5 alkynylhydroxyamine, C<sub>3</sub>-C<sub>50</sub> branched chain substituted  
alkynylhydroxyamine, C<sub>3</sub>-C<sub>50</sub> cyclic oxyalkynylamine, C<sub>3</sub>-C<sub>50</sub>  
cyclic substituted oxyalkynylamine, C<sub>1</sub>-C<sub>20</sub> straight chain  
alkylsemicarbazide, C<sub>1</sub>-C<sub>20</sub> straight chain substituted  
alkylsemicarbazide, C<sub>2</sub>-C<sub>50</sub> branched chain alkylsemicarbazide,  
10 C<sub>2</sub>-C<sub>50</sub> branched chain substituted alkylsemicarbazide, C<sub>3</sub>-C<sub>50</sub>  
cyclic alkylsemicarbazide, C<sub>3</sub>-C<sub>50</sub> cyclic substituted  
alkylsemicarbazide, C<sub>2</sub>-C<sub>20</sub> straight chain  
alkenylsemicarbazide, C<sub>2</sub>-C<sub>20</sub> straight chain substituted  
alkenylsemicarbazide, C<sub>3</sub>-C<sub>50</sub> branched chain  
15 alkenylsemicarbazide, C<sub>3</sub>-C<sub>50</sub> branched chain substituted  
alkenylsemicarbazide, C<sub>3</sub>-C<sub>50</sub> cyclic alkenylsemicarbazide, C<sub>3</sub>-  
C<sub>50</sub> cyclic substituted alkenylsemicarbazide, C<sub>2</sub>-C<sub>20</sub> straight  
chain alkynylsemicarbazide, C<sub>2</sub>-C<sub>20</sub> straight chain substituted  
alkynylsemicarbazide, C<sub>3</sub>-C<sub>50</sub> branched chain  
20 alkynylsemicarbazide, C<sub>3</sub>-C<sub>50</sub> branched chain substituted  
alkynylsemicarbazide, C<sub>3</sub>-C<sub>50</sub> cyclic alkynylsemicarbazide, C<sub>3</sub>-  
C<sub>50</sub> cyclic substituted alkynylsemicarbazide, C<sub>1</sub>-C<sub>20</sub> straight  
chain alkylthiosemicarbazide, C<sub>1</sub>-C<sub>20</sub> straight chain  
substituted alkylthiosemicarbazide, C<sub>2</sub>-C<sub>50</sub> branched chain  
25 alkylthiosemicarbazide, C<sub>2</sub>-C<sub>50</sub> branched chain substituted  
alkylthiosemicarbazide, C<sub>3</sub>-C<sub>50</sub> cyclic alkylthiosemicarbazide,  
C<sub>3</sub>-C<sub>50</sub> cyclic substituted alkylthiosemicarbazide, C<sub>2</sub>-C<sub>20</sub>  
straight chain alkenylthiosemicarbazide, C<sub>2</sub>-C<sub>20</sub> straight  
chain substituted alkenylthiosemicarbazide, C<sub>3</sub>-C<sub>50</sub> branched  
30 chain alkenylthiosemicarbazide, C<sub>3</sub>-C<sub>50</sub> branched chain  
substituted alkenylthiosemicarbazide, C<sub>3</sub>-C<sub>50</sub> cyclic  
alkenylthiosemicarbazide, C<sub>3</sub>-C<sub>50</sub> cyclic substituted  
alkenylthiosemicarbazide, C<sub>2</sub>-C<sub>20</sub> straight chain  
alkynylthiosemicarbazide, C<sub>2</sub>-C<sub>20</sub> straight chain substituted  
35 alkynylthiosemicarbazide, C<sub>3</sub>-C<sub>50</sub> branched chain  
alkynylthiosemicarbazide, C<sub>3</sub>-C<sub>50</sub> branched chain substituted  
alkynylthiosemicarbazide, C<sub>3</sub>-C<sub>50</sub> cyclic



alkynylthiosemicarbazide,  $C_3-C_{50}$  cyclic substituted  
 alkynylthiosemicarbazide,  $C_1-C_{20}$  straight chain  
 alkylhydrazone,  $C_1-C_{20}$  straight chain substituted  
 alkylhydrazone,  $C_2-C_{50}$  branched chain alkylhydrazone,  $C_2-C_{50}$   
 5 branched chain substituted alkylhydrazone,  $C_3-C_{50}$  cyclic  
 hydrazoalkane,  $C_3-C_{50}$  cyclic substituted hydrazoalkane,  $C_2-C_{20}$   
 straight chain alkenylhydrazone,  $C_2-C_{20}$  straight chain  
 substituted alkenylhydrazone,  $C_3-C_{50}$  branched chain  
 alkenylhydrazone,  $C_3-C_{50}$  branched chain substituted  
 10 alkenylhydrazone,  $C_3-C_{50}$  cyclic hydrazoalkene,  $C_3-C_{50}$  cyclic  
 substituted hydrazoalkene,  $C_2-C_{20}$  straight chain  
 alkynylhydrazone,  $C_2-C_{20}$  straight chain substituted  
 alkynylhydrazone,  $C_3-C_{50}$  branched chain alkynylhydrazone,  $C_3-$   
 $C_{50}$  branched chain substituted alkynylhydrazone,  $C_3-C_{50}$  cyclic  
 15 hydrazoalkyne,  $C_3-C_{50}$  cyclic substituted hydrazoalkyne,  $C_1-C_{20}$   
 straight chain alkylhydrazide,  $C_1-C_{20}$  straight chain  
 substituted alkylhydrazide,  $C_3-C_{50}$  branched chain  
 alkylhydrazide,  $C_3-C_{50}$  branched chain substituted  
 alkylhydrazide,  $C_3-C_{50}$  cyclic alkylhydrazide,  $C_3-C_{50}$  cyclic  
 20 substituted alkylhydrazide,  $C_2-C_{20}$  straight chain  
 alkenylhydrazide,  $C_2-C_{20}$  straight chain substituted  
 alkenylhydrazide,  $C_3-C_{50}$  branched chain alkenylhydrazide,  $C_3-$   
 $C_{50}$  branched chain substituted alkenylhydrazide,  $C_3-C_{50}$  cyclic  
 alkenylhydrazide,  $C_3-C_{50}$  cyclic substituted  
 25 alkenylhydrazide,  $C_2-C_{20}$  straight chain alkynylhydrazide,  $C_2-$   
 $C_{20}$  straight chain substituted alkynylhydrazide,  $C_3-C_{50}$   
 branched chain alkynylhydrazide,  $C_3-C_{50}$  branched chain  
 substituted alkynylhydrazide,  $C_3-C_{50}$  cyclic alkynylhydrazide  
 and  $C_3-C_{50}$  cyclic substituted alkynylhydrazide.

30 As noted above, the conjugate groups of the invention  
 may be attached to the 4'-position of the nucleosidic sugar  
 via a linker. Further linking groups include  $\Omega$ -aminoalkoxy  
 moieties and  $\Omega$ -aminoalkylamino moieties for linking a  
 conjugate group to a 4'-hydroxyl group. Many other linking  
 35 groups are known including many that are commercially  
 available, including heterobifunctional and homobifunctional  
 linking moieties available from the Pierce Co. (Rockford,

11).

A further preferred group of linking groups include those groups used as the internucleoside linkages between 4'-desmethyl nucleosides of the parent application, application serial number 039,834, filed March 30, 1993, the entire contents of which are herein expressed incorporated by reference.

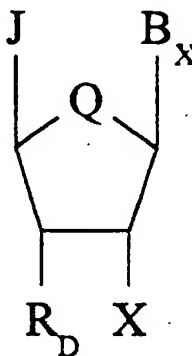
In one embodiment for attaching a conjugate molecule to a 4-desmethyl nucleoside, a linking moiety located on a 4-desmethyl nucleoside compound of the invention may be reacted with an active ester derivative of a conjugation moiety of the invention (e.g., cholic acid). Such linking moieties are particularly useful in extending polyalkylamine moieties extending from the oligomeric compounds of the invention. Active ester derivatives are well known to those skilled in the art. Representative active esters include N-hydroxysuccinimide esters, tetrafluorophenolic esters, pentafluorophenolic esters and pentachlorophenolic esters. For cholic acid, the reaction of a pendant amino group and the active ester produces an oligomer in which cholic acid is attached to the N-terminal position through a linking group. Cholic acid or other such pendent groups can be attached to the carboxy terminal of the oligomeric compound by conversion to the N-hydroxysuccinimide ester thereof, and then by further reaction with the oligomeric compound of interest.

Conjugate molecules of the invention can also be directly connected to the 4'-desmethyl nucleosides compounds of the invention without use of linkers. As for instance a conjugate molecule bearing a hydroxyl group, a thiol group, or an amino group is reacted with a 4'-desmethyl nucleoside compound of the invention that has a leaving group at the 4' position. The OH, SH or NH<sub>2</sub> moiety of the conjugate groups displaces the leaving group to directly join the conjugate to the 4'-desmethyl nucleoside compound of the invention.

The oligomeric compounds of the invention are

synthesized utilizing at least two different methods, both of which can use art recognized solid state synthesis or solution phase synthesis. In a first process, the "4-des-methyl" nucleoside that is to be functionalized with a conjugate group can be attached to a growing oligomeric compound and then so functionalized. In a second process the 4'-desmethyl nucleoside monomer can be first functionalized and then attached to the remainder of the oligomeric compound.

10 A conjugate group of a compound of the invention can be linked to a 4'-desmethyl nucleoside monomer or oligomer through displacement of a leaving group from the 4'-position (as numbered using the sugar ring numbering of a pentofuranosyl nucleoside) of a nucleoside, or the 4'-  
15 terminal position (as numbered using the sugar ring numbering of a pentofuranosyl nucleoside) of an oligonucleoside. In preferred embodiments, the 4'-desmethyl compound has structure:



where J is a leaving group. Leaving groups are functional moieties which are designed to be easily displaced. Representative leaving groups include alkyl and arylsulfonyls including p-toluenesulfonyl (tosyl), 4-dimethylaminoazobenzenesulfonyl (dabsyl), 5-dimethylaminonaphthalenesulfonyl (dansyl), trifluoromethylsulfonyl (triflate), methylsulfonyl (mesyl); halogens; o-trichloroacetimidates; 2,4,6-trichlorophenyl; dialkyl phosphite and acyloxy leaving groups including acetoxy, benzoyloxy, p-methoxybenzoyloxy and p-methylbenzoyloxy and

other known leaving groups. Acyloxy leaving groups ( $-OR_E$  where  $R_E$  is  $C(O)-$ ) are preferred, particularly  $OC(O)CH_3$ .

$R_D$  can be H, hydroxyl, a nucleoside, an activated nucleotide, a nucleotide, an oligonucleotide or an  
5 oligonucleoside or a protected derivative thereof.

$B_x$  is a naturally occurring or non-naturally occurring heterocyclic base as described *supra*, and also includes such bases bearing protecting groups.

The leaving group can be attached above or below the  
10 plane of the sugar ring, i.e., cis or trans to the heterocyclic base. The leaving group may therefore be displaced by either a bimolecular or unimolecular mechanism, dependent upon the specific leaving group and reaction conditions employed. In preferred embodiments the leaving  
15 group is trans to the heterocyclic base, and is displaced via a bimolecular nucleophilic displacement.

In preferred methods of the invention compounds having the above noted structure can be incorporated into a wide variety of oligomeric compounds by standard chemistries.  
20 Thus,  $R_D$  can be an activated phosphorous group. In the context of the invention activated phosphorous groups are groups which are used to facilitate the creation of internucleosidic bonds in nucleic acid synthetic regimes. Thus, activated phosphorous groups include phosphate groups  
25 having  $P^V$  such as, for example, activated phosphate esters, and phosphite groups having  $P^{III}$  such as phosphoramidites or H-phosphonates. An activated nucleotide according to the invention is a nucleoside which bears an activated phosphorous group.

30 In preferred processes of the invention a compound of the above structure is incorporated as the terminal nucleoside unit of a preformed oligomeric structure. Such oligomeric structures can be any of the various natural or synthetic oligonucleotides or oligonucleotide mimics as known  
35 in the art.

In accordance with the present invention, methods which are amenable to automated synthetic schemes,

especially solid-state support synthetic schemes, are preferred. While a number of methodologies can be employed, one preferred methodology follows. A nucleoside analog is attached to a solid support in any conventional fashion.

5 It is customary, and preferred, to employ a linker to a solid support such as a polymeric carrier at the 3' position. An oligonucleotide or oligonucleoside is then synthesized as per any of the various known solid state technologies. A final 4'-desmethyl nucleotide is then added

10 to the growing oligomeric compound.

In one process of the invention, the 4'-desmethyl nucleotide is substituted in the 4'-position with leaving group J as described above. The leaving group is then displaced with a linker or directly with a conjugate group

15 of interest. If a linker is utilized the conjugate group is then attached to the linker, either while still on the solid support or post removal from the solid support.

Alternately the 4'-desmethyl nucleotide is first reacted with a linker or directly with a conjugate group.

20 The modified nucleotide is then attached to the growing oligomeric compound on the solid support.

The 4'-desmethyl nucleoside (or nucleotide) is prepared with any base or base analog, B<sub>x</sub>, and either a pentofuranosyl moiety, where Q is oxygen, a cyclopentane moiety or fluorinated analogue thereof where Q is CH<sub>2</sub>, CHF or CF<sub>2</sub>, or a 4'-deoxy-4'-thio moiety, where Q is S. In certain preferred embodiments of the invention, the nucleoside is a ribonucleoside having a 2'-hydroxyl functionality. Such a functionality can be any of various

25 2'-groups known in the art. If the 2' functionality is OH (a normal ribonucleotide), the 2'-hydroxyl groups can be protected as necessary by means well known to persons of ordinary skill in the art. The only requirement for this group and for its protection is that the same be designed

30 so as to not interfere with substantive reactions in accordance with this invention. In other preferred embodiments, the 2'-hydroxyl group will be replaced with

35

other functional groups as for example 2'-alkoxy groups, 2'-alkoxy groups that are substituted with other groups including imidazole and other heterocyclic and amino groups, 2'-halogen, particularly fluoro. Particularly useful 2'-O-substituent groups include 2'-aminopropoxy, 2'-imidazolylbutoxy, 2'-methoxy, 2'-ethoxy, 2'-propoxy, 2'-methoxyethoxy and 2'-deoxy-2'-fluoro.

Additional objects, advantages, and novel features of this invention will become apparent to those skilled in the art upon examination of the following examples thereof, which are not intended to be limiting.

#### EXAMPLE 1

##### 5'-O-tert-Butyldiphenylsilylthymidine, 2.

A stirring solution of thymidine (50.0 g, 207 mmol) and DMAP (10 mg,  $8.2 \times 10^{-2}$  mmol) in 400 mL of pyridine was treated with TBDPSCl (43.4 g, 248 mmol) at 25°C for 16 h. The solvent was removed under reduced pressure and the residue was diluted with 1 L of AcOEt. The mixture was washed with 5% aqueous HCl (2 x 100 mL) and H<sub>2</sub>O (100 mL). The organic layer was dried (MgSO<sub>4</sub>) and concentrated under reduced pressure. The product was purified by silica gel chromatography (CH<sub>2</sub>Cl<sub>2</sub>/MeOH 20:1) to give 87.3 g (88%) of 2 as a white solid. An analytical sample was crystallized from diethylether. mp 164-166 °C (170-171 °C per Matulic-Adamic, J. Chem. Soc., Chem. Comm. 1985, 21, 1535) R<sub>f</sub>(CH<sub>2</sub>Cl<sub>2</sub>/MeOH 10:1) 0.31. <sup>1</sup>H-NMR (CDCl<sub>3</sub>): 1.08 (s, 9H, C-Me<sub>3</sub>), 1.61 (s, 3H, 5-Me), 2.18 (ddd, 1H, J=13.8, 8.5, 6.0 Hz, 2'-H<sub>a</sub>), 2.19 (br s, 1H, D<sub>2</sub>O exchangeable, 3'-OH), 2.44 (ddd, 1H, J=13.8, 5.6, 2.1 Hz, 2'-H<sub>b</sub>), 3.25 (br s, 1H, 5'-OH, D<sub>2</sub>O exchangeable), 3.85 (dd, 1H, J=11.5, 2.5 Hz, 5'-CHH), 3.98 (dd, 1H J=11.5, 2.3 Hz, 5'-CHH), 4.06 (dd, 1H, J=2.5, 2.3 Hz, 4'-H), 4.55 (dd, 1H J=6.0, 5.6 Hz, 3'-H), 6.43 (dd, 1H, J=8.5, 5.6 Hz, 1'-H), 7.26-7.51 (m, 6H, aromatic-H), 7.64-7.68 (m, 5H, 6-H and aromatic-H), 9.57 (s, 1H, NH, D<sub>2</sub>O exchangeable).

**EXAMPLE 2**

**N<sup>3</sup>-Benzyloxymethyl-5'-O-tert-butyldiphenylsilylthymidine, 3.**

To a stirred solution of 2 (117.0 g, 243.8 mmol) and Hunig's base (diisopropylethylamine, 63.0 g, 487.5 mmol) in CH<sub>2</sub>Cl<sub>2</sub> (400 mL) at 23 °C was added a solution of benzyl chloromethyl ether (40.7 g, 260.0 mmol) over a 15 min. period. The resultant mixture was maintained at 23 °C and stirred for 14 h. Ether (1 L) was added to the mixture and the ethereal solution was washed with 10% aqueous HCl (2 x 100 mL) and H<sub>2</sub>O (200 mL). The organic layer was dried (MgSO<sub>4</sub>) and concentrated under reduced pressure. The residue was purified by silica gel column chromatography (CH<sub>2</sub>Cl<sub>2</sub>/AcOEt 40:1 then 10:1) to yield 128.9 g (88%) of 3 as a white solid. R<sub>f</sub>(CH<sub>2</sub>Cl<sub>2</sub>/AcOEt 10:1) 0.31. <sup>1</sup>H-NMR (CDCl<sub>3</sub>) : 1.10 (s, 9H, C-Me<sub>3</sub>), 1.65 (s, 3H, 5-Me), 2.16 (m, 1H, 2'-H<sub>a</sub>), 2.41 (m, 1H, 2'-H<sub>b</sub>), 2.53 (br s, 1H 3'-OH), 3.84 (d, 1H, J=8.8 Hz, 5'-CHH), 3.98 (d, 1H, J=8.8 Hz, 5'-CHH), 4.01 (s, 1H, 4'-H), 6.64 (br s, 1H, 3'-H), 4.70 (s, 2H, OCH<sub>2</sub>Ph), 5.50 (s, 2H, NCH<sub>2</sub>O), 6.41 (dd, 1H, J=8.0, 5.8 Hz, 1'-H), 7.20-7.50 (m, 13H, aromatic-H), 7.65-7.69 (m, 3H, 6-H and aromatic-H). <sup>13</sup>C-NMR (CDCl<sub>3</sub>): 12.86 (-5-CH<sub>3</sub>), 19.45 (+, C-Me<sub>3</sub>), 27.09 (-, C-Me<sub>3</sub>), 41.16 (+, 2'-C), 64.25 (+, 5'-C), 70.70 (+, O-C-Ph), 71.97 (-, 4'-C), 72.29 (+, N-C-O), 85.51 (-, 3'-C), 87.20 (-, 1'-C), 110.47 (+, 5-C), 127.72, 128.08, 128.36 (-, aromatic-C), 130.25 (-, 6-C), 132.42, 133.00 (+, aromatic-C) 134.39, 135.37, 135.62 (-, aromatic-C), 137.95 (+, aromatic-C), 151.01 (+, 2-C), 163.68 (+, 4-C).

**EXAMPLE 3**

**N<sup>3</sup>-Benzyloxymethyl-3'-O-benzoyl-5'-O-tert-butyldiphenylsilyl-thymidine, 4.**

A stirred solution of 3 (128.0 g, 213.3 mmol) in a 4:1 mixture of CH<sub>2</sub>Cl<sub>2</sub>/Et<sub>3</sub>N (500 mL) was treated with (48.4 g, 40 mL, 344.6 mmol) of BzCl at 23 °C for 8 h. The resultant precipitate was removed by filtration. The filtrate was concentrated under reduced pressure to leave the crude product as a brownish syrup. Purification of the syrup by

silica gel column chromatography (hexanes/AcOEt 10:1 then 1:1) gave 130.7 g (87%) of **4** as a white solid.  $R_f$  (Hexanes/AcOEt 1:1) 0.82.  $^1\text{H-NMR}$  ( $\text{CDCl}_3$ ): 1.40 (s, 9H, C-Me<sub>3</sub>), 1.60 (s, 3H, 5-Me), 2.37 (ddd, 1H,  $J=13.8$ , 9.3, 7.2 Hz, 2'-H<sub>g</sub>), 2.62 (dd, 1H,  $J=13.8$ , 4.3 Hz, 2'-H<sub>g</sub>), 4.09 (m, 2H, 5'-H), 4.26 (m, 1H, 4'-H), 4.74 (s, 2H, O-CH<sub>2</sub>-Ph), 5.54 (s, 2H, N-CH<sub>2</sub>-O), 5.71 (d, 1H,  $J=7.2$  Hz, 3'-H), 6.57 (dd, 1H,  $J=9.3$ , 4.3 Hz, 1'-H), 7.24-7.74 (m, 13H, aromatic-H), 8.05-8.15 (m, 3H, 6-H and aromatic-H).  $^{13}\text{C-NMR}$  ( $\text{CDCl}_3$ ): 12.82 (-, 5-Me), 19.54 (+, C-Me<sub>3</sub>), 27.16 (C-Me<sub>3</sub>), 38.28 (+, 2'-C), 64.41 (+, 5'-C), 70.80 (+, O-C-Ph), 72.29 (+, N-C-O), 75.60 (-, 4'-C), 85.28 (-, 1'-C and 3'-C), 110.95 (+, 5-C), 127.72, 128.23, 128.37, 128.50, 128.63 (-, aromatic-C), 129.43 (+, aromatic-C), 129.84, 130.22 (-, aromatic-C), 132.14, 133.07 (+, aromatic-C), 133.60 (-, 6-C), 133.94, 135.32, 135.65 (-, aromatic-C), 138.15 (+, aromatic-C), 151.19 (+, 2-C), 163.50 (+, 4-C), 166.11 (+, benzoyl C=O).

#### EXAMPLE 4

##### N<sup>3</sup>-Benzyloxymethyl-3'-O-benzoylthymidine, **5**.

The silyl ether **4** (96.0 g, 136.4 mmol) in THF (600 mL) was treated with hydrogen fluoride-pyridine (70% HF in pyridine, 30 mL) at 0 °C for 4 h under a N<sub>2</sub> atmosphere. The resultant mixture was diluted with AcOEt (600 mL) and washed with H<sub>2</sub>O (2 x 300 mL). The organic layer was dried (MgSO<sub>4</sub>) and concentrated at reduced pressure. The residue was purified by silica gel column chromatography (CH<sub>2</sub>Cl<sub>2</sub>/AcOEt 10:1) to give 61.6 g (97%) of **5** as a white solid.  $R_f$  (CH<sub>2</sub>Cl<sub>2</sub>/AcOEt 10:1) 0.29.  $^1\text{H-NMR}$  ( $\text{CDCl}_3+\text{D}_2\text{O}$ ): 1.95 (s, 3H, 5-Me), 2.53 (m, 2H, 2'-H), 4.00 (m, 2H, 5'-H), 4.25 (m, 1H, 4'-H), 4.71 (s, 2H, O-CH<sub>2</sub>-Ph), 5.51 (s, 2H, N-CH<sub>2</sub>-O), 5.60 (m, 1H 3'-H), 6.36 (dd, 1H,  $J=7.6$ , 6.6 Hz, 1'-H), 7.25-7.66 (m, 9H, 6-H and aromatic-H), 8.05 (d, 2H,  $J=7.1$  Hz, aromatic-H).  $^{13}\text{C-NMR}$  ( $\text{CDCl}_3$ ): 13.29 (-, 5-Me), 37.82 (+, 2'-C), 62.54 (+, 5'-C), 70.73 (+, O-C-Ph), 72.25 (+, N-C-O), 75.82 (-, 4'-C), 85.43 (-, 3'-C), 86.13 (-, 1'-C), 110.41 (+, 5-C), 127.65, 128.36, 128.59 (-, aromatic-C), 129.34 (+,



aromatic-C), 129.69 (-, 6-C), 133.60, 135.40 (-, aromatic-C), 137.87 (+, aromatic-C), 151.18 (+, 2-C), 163.65 (+, 4-C), 166.11 (+, benzoyl C=O).

#### EXAMPLE 5

5    **tert-Butyl    N<sup>3</sup>-Benzoxymethyl-3'-O-benzoylthymidine-5'-carboxylate, 6**

The reaction was performed as described by Corey and Samuelsson, *J. Org. Chem.* 1984, 49, 4735, for the oxidation of 5'-OH of the uridine derivative to its corresponding 5'-  
10    tert-butyl carboxylate. Chromium(VI) oxide (31.4 g, 314.2 mmol) in CH<sub>2</sub>Cl<sub>2</sub> (480 mL) was cooled to 0 °C and then pyridine (49.7 g, 638.4 mmol) in DMF (120 mL) was added dropwise to the reaction mixture (caution: extremely exothermic) over a period of 1 h. The mixture was stirred  
15    at 0 °C for 30 min. Alcohol 5 (36.6 g, 78.5 mmol) in CH<sub>2</sub>Cl<sub>2</sub>/DMF (4:1 v/v, 100 mL) was added followed by acetic anhydride (64.1 g, 628.4 mmol) and t-BuOH (116.4 g, 1.57 mmol). The resultant mixture was warmed to 23 °C and stirred for 18 h. Ethanol (20 mL) was added to the reaction  
20    and the mixture was stirred for additional 15 min. The reaction mixture was poured into AcOEt (400 mL) and the insoluble material was filtrated through a Buchner funnel padded with 200 g of silica gel and 50 g of MgSO<sub>4</sub>. The solid left in the funnel was rinsed with AcOEt (4 x 100 mL).  
25    The combined filtrate and rinses were concentrated under reduced pressure. The residue was purified by silica gel column chromatography (hexanes/AcOEt 3:1) to give 25.6 g (61%) of 5 as a white solid. An analytical sample (about 400 mg) was crystallized from ether/hexanes afforded white,  
30    needle-like crystals. mp 80-82°C. R<sub>f</sub>(hexanes/AcOEt 3:1) 0.23. <sup>1</sup>H-NMR(CDCl<sub>3</sub>): 1.55 (s, 9H, C-Me<sub>3</sub>), 2.21 (ddd, 1H, J=14.3, 9.1, 5.0 Hz, 2'-H<sub>a</sub>), 2.61 (dd, 1H, J=14.3, 5.2 Hz, 2'-H<sub>b</sub>), 4.63 (s, 1H, 4'-H), 4.71 (s, 2H, O-CH<sub>2</sub>-Ph), 5.51 (s, 2H, N-CH<sub>2</sub>-O), 5.65 (d, 1H, J=5.0 Hz, 3'-H), 6.61 (dd, 1H, J=9.1, 5.2 Hz, 1'-H), 7.24-7.63 (m, 8H, aromatic-H), 8.07  
35    (d, 2H, J=7.1 Hz, aromatic-H), 8.09 (s, 1H, 6-H). <sup>13</sup>C-

NMR(CDCl<sub>3</sub>) : 13.40 (-, 5-Me), 27.92 (-, C-Me<sub>3</sub>), 36.65 (+, 2'-C), 70.58 (+, O-C-Ph), 72.09 (+, N-C-O), 76.66 (-, 4'-C), 82.55 (-, 3'-C), 83.36 (+, C-Me<sub>3</sub>), 86.88 (-, 1'-C), 110.61 (+, 5-C), 127.55, 127.24 128.23, 128.53 (-, aromatic-C), 128.99 (+, aromatic-C), 129.78 (-, 6-C), 133.71, 134.72 (-, aromatic-C), 138.06 (+, aromatic-C), 151.14 (+, 2-C), 163.28 (+, 4-C), 165.26 (+, benzoyl C=O), 169.40 (+, 5'-C).

**EXAMPLE 6**

**N<sup>3</sup>-Benzoxymethyl-3'-O-benzoylthymidine-5'-carboxylic acid,**

10 **7.**

A solution of 6 (22.0 g, 41.0 mmol) in CF<sub>3</sub>COOH (100 mL) was stirred at 23 °C for 2 h. Toluene (200 mL) was then added and the mixture was concentrated under reduced pressure. The coevaporation of toluene was repeated twice to ensure complete removal of the CF<sub>3</sub>COOH. The resultant light yellow powder was purified by silica gel column chromatography (CH<sub>2</sub>Cl<sub>2</sub>/AcOEt 8:1) to afford 19.3 g (87%) of 7 as a white powder. R<sub>f</sub>(CHCl<sub>3</sub>/MeOH 4:1) 0.39, <sup>1</sup>H-NMR(CDCl<sub>3</sub>+D<sub>2</sub>O): 1.95 (s, 3H, 5-Me), 2.27 (ddd, 1H, J=14.3, 9.1, 4.9 Hz, 2'-H<sub>a</sub>), 2.68 (dd, 1H, J=14.3, 5.2 Hz, 2'-H<sub>b</sub>), 4.71 (s, 2H, O-CH<sub>2</sub>-Ph), 4.79 (s, 1H, 4'-H), 5.52 (s, 2H, N-CH<sub>2</sub>-O), 5.76 (d, 1H, J=4.9 Hz, 3'-H), 6.55 (dd, 1H, J=9.1, 5.2 Hz, 1'-H), 7.24-7.60 (m, 8H, aromatic-H), 7.97 (s, 1H, 6-H), 8.06 (d, 2H, J=7.1 Hz, aromatic-H). <sup>13</sup>C-NMR(CDCl<sub>3</sub>): 13.42 (+, 5-Me), 36.68 (+, 2'-C), 70.83 (+, O-C-Ph), 73.38 (+, N-C-O), 76.93 (-, 4'-C), 82.01 (-, 3'-C), 87.58 (-, 1'-C), 110.80 (+, 5-C), 127.72, 127.86, 128.39, 128.70 (-, aromatic-C), 128.71 (+, aromatic-C), 129.90 (-, 6-C), 133.95, 135.63 (-, aromatic-C), 128.53 (+, aromatic-C), 11.20 (+, 2-C), 163.94 (+, 4-C), 165.61 (+, benzoyl C=O), 171.93 (+, 5'-C).

**EXAMPLE 7**

**(2'S, 4'S, 5'R)-1-[5'-Acetoxy-4'-benzoyloxy-tetrahydrofuran-2'-yl]-N<sup>3</sup>-benzoxymethylthymine, 8.**

35 A stirred solution of 7 (10.6 g, 22.0 mmol) in DMF (75

mL) was treated with  $\text{Pb}(\text{OAc})_4$  (11.8 g, 26.5 mmol) at 23 °C for 2 h under darkness. The mixture was diluted with AcOEt (250 mL) and the resultant suspension was filtrated through a Celite pad (50 g). The solid was rinsed several times  
5 with AcOEt. The combined filtrate and rinses were concentrated under reduced pressure. The residue was purified by silica gel column chromatography (hexanes/AcOEt 1:1) to give 6.5 g (60%) of a 3:7  $\alpha/\beta$  (as determined by the  $^1\text{H}$ -NMR 2'-H ratio) anomeric mixture of **8** as a light yellow  
10 syrup. An aliquot of the anomeric mixture (~0.2 g) was separated on a silica gel column (hexanes/AcOEt 8:1  $\rightarrow$  2:1 gradient) to afford 53 mg of **8 $\alpha$**  and 121 mg of **8 $\beta$** , both as white foams.

**8 $\alpha$** :  $R_f$ (hexanes/AcOEt, 1:1) 0.53.  $^1\text{H}$ -NMR ( $\text{CDCl}_3$ ): 1.95 (d, 3H,  $J$  = 1.1 Hz, 5- $\text{CH}_3$ ), 2.05 (s, 3H, acetoxy- $\text{CH}_3$ ), 2.52-2.75 (m, 2H, 3'-H), 4.70 (s, 2H,  $\text{OCH}_2\text{Ph}$ ), 5.49 (s, 2H,  $\text{NCH}_2\text{O}$ ), 5.60-5.65 (m, 1H, 4'-H), 6.40 (dd, 1H,  $J$  = 7.9, 3.8 Hz, 2'-H), 6.72 (d, 1H,  $J$  = 4.0 Hz, 5'-H), 7.06 (d, 1H,  $J$  = 1.1 Hz, 6-H), 7.22-7.60 (m, 8H, aromatic-H), 8.02 (d, 2H,  $J$  = 7.6  
20 Hz, aromatic-H).  $^{13}\text{C}$ -NMR ( $\text{CDCl}_3$ ): 13.04 (-, 5- $\text{CH}_3$ ), 20.73 (-, acetoxy- $\text{CH}_3$ ), 33.81 (+, 3'-C), 70.45 (+, O-C-Ph), 71.08 (-, 4'-C), 72.11 (+, N-C-O), 85.47 (-, 2'-C), 94.10 (-, 5'-C), 110.90 (+, 5-C), 127.47, 128.11, 128.45 (-, aromatic-C), 128.64 (+, aromatic-C), 129.54 (-, 6-C), 133.55, 133.92 (-, aromatic-C), 137.75 (+, aromatic-C), 150.55 (+, 2-C), 163.01 (+, 4-C), 165.43 (+, benzoyl C=O), 169.18 (+, acetoxy C=O).  
25 Anal. Calcd. for  $\text{C}_{26}\text{H}_{26}\text{N}_2\text{O}_8$ : C, 63.16; H, 5.26; N, 5.67. Found: C, 63.12; H, 5.38; N, 5.45.

**8 $\beta$** :  $R_f$ (hexanes/AcOEt 1:1) 0.47.  $^1\text{H}$ -NMR ( $\text{CDCl}_3$ ): 1.95 (s, 3H, 5- $\text{CH}_3$ ), 2.18 (s, 3H, acetoxy- $\text{CH}_3$ ), 2.33 (ddd, 1H,  $J$  = 14.9, 8.2, 4.9 Hz, 3'- $\text{H}_\text{a}$ ), 2.75 (dd, 1H,  $J$  = 14.9, 6.2 Hz, 3'- $\text{H}_\text{b}$ ), 4.70 (s, 2H,  $\text{OCH}_2\text{Ph}$ ), 5.50 (s, 2H,  $\text{NCH}_2\text{O}$ ), 5.52 (d, 1H,  $J$  = 4.9 Hz, 4'-H), 6.42 (s, 1H, 5'-H), 6.73 (dd, 1H,  $J$  = 8.2, 6.2 Hz, 2'-H), 7.22-7.63 (m, 9H, 6-H and aromatic-H), 8.04  
35 (d, 2H,  $J$  = 7.5 Hz, aromatic-H).  $^{13}\text{C}$ -NMR ( $\text{CDCl}_3$ ): 13.40 (-, 5- $\text{CH}_3$ ), 20.82 (-, acetoxy- $\text{CH}_3$ ), 34.95 (+, 3'-C), 70.51 (+, O-C-Ph), 72.07 (+, N-C-O), 76.64 (-, 4'-C), 87.15 (-, 2'-C),

98.61 (-, 5'-C), 111.03 (+, 5-C), 127.45, 128.11, 128.45 (-, aromatic-C), 129.68 (-, 6-C), 133.02, 133.69 (-, aromatic-C), 137.77 (+, aromatic-C), 150.91 (+, 2-C), 162.84 (+, 4-C), 165.12 (+, benzoyl C=O), 169.13 (+, acetoxy C=O). Anal.  
5 Calcd for  $C_{26}H_{26}N_2O_8 \cdot H_2O$ : C, 60.94; H, 5.47; N, 5.47. Found: C, 60.98; H, 5.18; N, 5.30.

**EXAMPLE 8**

(2'R,4'S,5'S)-(4'-Benzoxo-5'-O-polyethyleneglycoltetrahydrofuran-2'-yl)-N<sup>3</sup>-(benzyloxymethyl)thymine (9)

10 A solution of 3.7 g (7.5 mmol) of acetate 8, 25-35 mmol of PEG alcohol, and 4 mL (3.0 g, 30 mmol) of Et<sub>3</sub>N in 150 mL of CH<sub>2</sub>Cl<sub>2</sub> was added 14 mL (16.2 g, 75 mmol) of TMSOTf at -23 °C under Ar atmosphere. The reaction mixture was stirred at -23 °C for 2 h, then placed in the freezer (-15  
15 °C) for 16-24 h. The resultant mixture was poured into a 500 mL bilayer mixture of AcOEt/H<sub>2</sub>O (9:1) containing 15 mL of Et<sub>3</sub>N. The organic layer was dried (MgSO<sub>4</sub>) and concentrated at reduced pressure. The residue was purified by SiO<sub>2</sub> column chromatography to yield 9.

20 9a: yield 84.9% (colorless syrup). R<sub>f</sub>(hexanes/AcOEt 1:1) 0.22. <sup>1</sup>H NMR (CDCl<sub>3</sub>+D<sub>2</sub>O) δ 1.95 (s, 3H, 5-CH<sub>3</sub>), 2.34 (ddd, J = 14.3, 8.1, 5.1 Hz, 3'-H<sub>a</sub>), 2.62 (dd, 1H, J = 14.3, 6.5 Hz, 3'-H<sub>b</sub>), 3.60-3.95 (m, 4H), 4.70 (s, 2H, OCH<sub>2</sub>Ph), 5.25 (s, 1H, 5'-H), 5.47 (d, J = 5.1 Hz, 4'-H),  
25 5.50 (s, 2H, NCH<sub>2</sub>O), 6.80 (dd, 1H, J = 8.1, 6.5 Hz, 2'-H), 7.18-7.65 (m, 9H, 6-H and aromatic-H), 8.04-8.07 (m, 2H, aromatic-H). <sup>13</sup>C NMR (CDCl<sub>3</sub>+D<sub>2</sub>O) δ 13.16, 35.00 (3'-C), 61.41, 70.33, 70.76, 72.29, 77.67 (4'-C), 86.53 (2'-C), 106.61 (5'-C), 111.56 (5-C), 127.66, 128.06, 128.33 (6-C),  
30 128.97, 129.84, 133.80, 134.38, 137.99, 151.41 (2-C), 163.35 (4-C), 165.76 (benzoyl C=O).

9b: yield 84.3% (colorless syrup). R<sub>f</sub>(hexanes/AcOEt 1:1) 0.18. <sup>1</sup>H NMR (CDCl<sub>3</sub>) δ 1.95 (s, 3H, 5-CH<sub>3</sub>), 2.34 (ddd, 1H, J = 14.7, 8.1, 4.9 Hz, 3'-H<sub>a</sub>), 2.62 (dd, 1H, J = 14.7,  
35 6.3 Hz, 3'-H<sub>b</sub>), 3.36 (s, 3H, CH<sub>3</sub>O), 3.54-3.98 (m, 12 H), 4.72 (s, 2H, OCH<sub>2</sub>Ph), 5.24 (s, 1H, 5'-H), 5.49 (d, 1H, J =

4.9 Hz, 4'-H), 5.52 (s, 2H, NCH<sub>2</sub>O), 6.80 (dd, 1H, *J* = 8.1, 6.3 Hz, 2'-H), 7.26-7.65 (m, 9H, 6-H and aromatic-H), 8.02-8.06 (m, 2H, aromatic-H). <sup>13</sup>C NMR (CDCl<sub>3</sub>) δ 13.29, 35.06 (3'-C), 59.04 (CH<sub>3</sub>O), 68.00, 70.14, 70.61, 71.91, 72.19, 77.51 (4'-C), 86.38 (2'-C), 106.43 (5'-C), 111.43 (5-C), 127.65, 128.31, 128.60 (6-C), 129.07, 129.81, 133.70, 134.45, 138.04, 151.41 (2-C), 163.28 (4-C), 165.51 (benzoyl C=O).

9c: 76.5% yield (white foam). *R<sub>f</sub>*(hexanes/AcOEt 1:1) 0.05. <sup>1</sup>H NMR (CDCl<sub>3</sub>) δ 1.98 (s, 3H, 5-CH<sub>3</sub>), 2.37 (ddd, 1H, *J* = 14.7, 8.2, 4.9 Hz, 3'-H<sub>a</sub>), 2.62 (dd, 1H, *J* = 14.7, 6.5 Hz, 3'-H<sub>b</sub>), 2.78 (br s, 1H, OH, D<sub>2</sub>O exchangeable), 3.57-3.99 (m, 16 H), 4.72 (s, 2H, OCH<sub>2</sub>Ph), 5.25 (s, 1H, 5'-H), 5.50 (d, 1H, *J* = 4.9 Hz, 4'-H), 5.52 (s, 2H, NCH<sub>2</sub>O), 6.80 (dd, 1H, *J* = 8.2, 6.5 Hz, 2'-H), 7.26-7.65 (m, 9H, 6-H and aromatic-H), 8.02-8.06 (m, 2H, aromatic-H). <sup>13</sup>C NMR (CDCl<sub>3</sub>) δ 13.27, 35.02 (3'-C), 61.70, 68.02, 70.15, 70.28, 70.54, 70.73, 72.26, 72.61, 77.61 (4'-C), 86.47 (2'-C), 106.55 (5'-C), 111.40 (5-C), 127.66, 128.31, 128.60 (6-C), 129.12, 129.84, 133.71, 134.48, 138.04, 151.45 (2-C), 163.35 (4-C), 165.61 (benzoyl C=O).

9d: 75.9% yield (yellowish syrup). *R<sub>f</sub>*(hexanes/AcOEt 1:1) 0.10. <sup>1</sup>H NMR (CDCl<sub>3</sub>) δ 1.96 (s, 3H, 5-CH<sub>3</sub>), 2.34 (ddd, 1H, *J* = 14.7, 7.9, 4.6 Hz, 3'-H<sub>a</sub>), 2.37 (br s, 1H, OH, D<sub>2</sub>O exchangeable), 2.62 (dd, 1H, *J* = 14.7, 6.8 Hz, 3'-H<sub>b</sub>), 3.44-4.01 (m, 24 H), 4.72 (s, 2H, OCH<sub>2</sub>Ph), 5.25 (s, 1H, 5'-H), 5.48 (d, 1H, *J* = 4.6 Hz, 4'-H), 5.52 (s, 2H, NCH<sub>2</sub>O), 6.79 (dd, 1H, *J* = 7.9, 6.8 Hz, 2'-H), 7.25-7.65 (m, 9H, 6-H and aromatic-H), 7.98-8.06 (m, 2H, aromatic-H). <sup>13</sup>C NMR (CDCl<sub>3</sub>) δ 13.03, 34.85 (3'-C), 61.28, 67.83, 70.12, 70.33, 71.38, 71.96, 72.47, 77.44 (4'-C), 86.22 (2'-C), 106.27 (5'-C), 111.08 (5-C), 127.39, 128.08, 128.44 (6-C), 128.94, 129.60, 133.54, 134.38, 137.92, 151.17 (2-C), 163.08 (4-C), 165.29 (benzoyl C=O).

#### EXAMPLE 9

(2'R,4'S,5'S)-(4'-Hydroxy-5'-O-polyethyleneglycoltetrahydrofuran-2'-yl)thymine (10)

A suspension of 4.5 mmol of 9 and 2-3 g of  $\text{Pd}(\text{OH})_2/\text{C}$  in 40 mL of 3:1 MeOH/acetone was stirred at 23 °C under  $\text{H}_2$  atmosphere for 16 h. The suspended material was filtrated through a pad of Celite in a Büchner funnel and the solid in the funnel was rinsed several times with acetone. The combined filtrate and rinses were concentrated at reduced pressure to give a light yellow syrup. The syrup was redissolved in 25 mL of MeOH and 0.3 g (7.5 mmol) of NaOH was added. The mixture was stirred at 23 °C for 6 h and concentrated at reduced pressure. The residue was purified by  $\text{SiO}_2$  column chromatography to yield 10.

10a: yield 88.3% (white solid). mp 210 °C (dec.).  $R_f(\text{CHCl}_3/\text{MeOH } 9:1)$  0.22.  $^1\text{H NMR}$  ( $\text{D}_2\text{O}$ )  $\delta$  1.84 (s, 3H, 5- $\text{CH}_3$ ), 2.34 (m, 2H, 3'-H), 3.57-3.83 (m, 4H), 4.42 (dd,  $J = 3.7$ , 2.1 Hz, 4'-H), 5.08 (s, 1H, 5'-H), 6.56 (t, 1H,  $J = 7.1$  Hz, 2'-H), 7.56 (s, 1H, 6-H).  $^{13}\text{C NMR}$  ( $\text{DMSO}-d_6$ )  $\delta$  12.07, 36.94 (3'-C), 59.98, 69.33, 74.11 (4'-C), 84.99 (2'-C), 108.50 (5'-C), 110.43 (5-C), 135.96 (6-C), 150.71 (2-C), 163.67 (4-C). Anal Calcd for  $\text{C}_{11}\text{H}_{16}\text{N}_2\text{O}_6 \cdot 0.7\text{CH}_3\text{OH}$ : C, 47.69; H, 6.38; N, 9.51. Found: C, 47.70; H, 6.01; N, 9.52.

10b: yield 90.6% (white solid). mp 54-56 °C.  $R_f(\text{CHCl}_3/\text{MeOH } 9:1)$  0.41.  $^1\text{H NMR}$  ( $\text{CDCl}_3$ )  $\delta$  1.93 (s, 3H, 5- $\text{CH}_3$ ), 2.11 (ddd, 1H,  $J = 14.2$ , 7.9, 4.9 Hz, 3'- $\text{H}_a$ ), 2.33 (br s, 1H, 4'-OH), 2.41 (dd, 1H,  $J = 14.2$ , 6.5 Hz, 3'- $\text{H}_b$ ), 3.38 (s, 3H,  $\text{CH}_3\text{O}$ ), 3.54-3.94 (m, 12 H), 4.37 (d, 1H,  $J = 4.9$  Hz, 4'-H), 5.14 (s, 1H, 5'-H), 6.68 (dd, 1H,  $J = 7.9$ , 6.5 Hz, 2'-H), 7.45 (s, 1H, 6-H), 9.25 (s, 1H, NH).  $^{13}\text{C NMR}$  ( $\text{CDCl}_3$ )  $\delta$  12.50, 37.42 (3'-C), 58.84 ( $\text{CH}_3\text{O}$ ), 67.52, 70.25, 70.40, 71.71, 75.01 (4'-C), 85.90 (2'-C), 109.23 (5'-C), 111.60 (5-C), 136.29 (6-C), 151.10 (2-C), 164.40 (4-C). Anal Calcd for  $\text{C}_{16}\text{H}_{26}\text{N}_2\text{O}_8 \cdot 2\text{H}_2\text{O}$ : C, 46.83; H, 7.32; N, 6.83. Found: C, 46.94; H, 7.50; N, 6.74.

10c: yield 85.9% (white foam).  $R_f(\text{CHCl}_3/\text{MeOH } 9:1)$  0.36.  $^1\text{H NMR}$  ( $\text{CDCl}_3$ )  $\delta$  1.92 (s, 3H, 5- $\text{CH}_3$ ), 2.09 (ddd, 1H,  $J = 14.2$ , 7.8, 4.9 Hz, 3'- $\text{H}_a$ ), 2.41 (dd, 1H,  $J = 14.2$ , 6.6 Hz, 3'- $\text{H}_b$ ), 3.58-3.90 (m, 16 H), 4.36 (d, 1H,  $J = 4.9$  Hz, 4'-H), 5.20 (s, 1H, 5'-H), 6.63 (dd, 1H,  $J = 7.8$ , 6.6 Hz,

2'-H), 7.46 (s, 1H, 6-H), 8.87 (s, 1H, NH).  $^{13}\text{C}$  NMR ( $\text{CDCl}_3$ )  $\delta$  12.53, 37.35 (3'-C), 61.26, 67.47, 70.00, 70.35, 72.51, 74.99, 77.71 (4'-C), 85.90 (2'-C), 109.27 (5'-C), 111.47 (5-C), 136.33 (6-C), 151.05 (2-C), 164.45 (4-C). Anal Calcd for  $\text{C}_{17}\text{H}_{28}\text{N}_2\text{O}_9 \cdot 0.5 \text{ H}_2\text{O}$ : C, 49.39; H, 7.02; N, 6.78. Found: C, 49.58; H, 7.12; N, 6.66.

10d: yield 87.3% (white wax).  $R_f$  ( $\text{CHCl}_3/\text{MeOH}$  10:1) 0.40.  $^1\text{H}$  NMR ( $\text{CDCl}_3$ )  $\delta$  1.92 (s, 3H, 5- $\text{CH}_3$ ), 2.08 (ddd, 1H,  $J$  = 13.8, 7.7, 4.8 Hz, 3'- $\text{H}_a$ ), 2.39 (dd, 1H,  $J$  = 13.8, 6.5 Hz, 3'- $\text{H}_b$ ), 2.81 (s, 1H, 4'-OH), 3.32-3.92 (m, 24 H), 4.01 (s, 1H, OH), 4.37 (d, 1H,  $J$  = 4.8 Hz, 4'-H), 5.16 (s, 1H, 5'-H), 6.66 (dd, 1H,  $J$  = 7.8, 6.5 Hz, 2'-H), 7.45 (s, 1H, 6-H), 8.90 (s, 1H, NH).  $^{13}\text{C}$  NMR ( $\text{CDCl}_3$ )  $\delta$  12.49, 37.50 (3'-C), 61.40, 67.53, 70.14, 70.41, 72.53, 75.01 (4'-C), 85.92 (2'-C), 109.27 (5'-C), 111.50 (5-C), 136.20 (6-C), 151.00 (2-C), 164.25 (4-C). Anal Calcd for  $\text{C}_{21}\text{H}_{36}\text{N}_2\text{O}_{11} \cdot 0.75 \text{ H}_2\text{O}$ : C, 49.85; H, 7.42; N, 5.54. Found: C, 49.79; H, 7.50; N, 5.60.

#### EXAMPLE 10

(2'R,4'S,5'S)-(4'-(2-Cyanoethyl-N,N-diisopropylaminophosphorityl)-5'-O-(dimethoxytrityloxytetraethyleneglycol)tetrahydrofuran-2'-yl)thymine (12)

A solution of 1.3 g (3.2 mmol) of 10c,  $1.0 \times 10^{-2}$  g ( $8.2 \times 10^{-2}$  mmol) of DMAP, and 1.0 g (10.0 mmol) of  $\text{Et}_3\text{N}$  in 30 mL of pyridine was treated with 1.6 g (4.8 mmol) of DMTr-Cl at 23 °C for 14 h. The solvent was removed at reduced pressure and the residue was purified by  $\text{SiO}_2$  column chromatography ( $\text{CHCl}_3/\text{MeOH}$  20:1 then 10:1) gave 1.9 g (83.6%) of 11 as a yellow foam.  $R_f$  ( $\text{CHCl}_3/\text{MeOH}$  20:1) 0.23.  $^1\text{H}$  NMR ( $\text{CDCl}_3$ )  $\delta$  1.92 (s, 3H, 5- $\text{CH}_3$ ), 2.04 (ddd, 1H,  $J$  = 14.2, 7.6, 4.8 Hz, 3'- $\text{H}_a$ ), 2.32 (dd, 1H,  $J$  = 14.2, 7.0 Hz, 3'- $\text{H}_b$ ), 3.23 (t, 2H,  $J$  = 5.0 Hz), 3.29 (br s, 1H, 4'-OH), 3.59-3.89 (m, 14 H), 3.79 (s, 6H, 2 x  $\text{OCH}_3$ ), 4.31 (d, 1H,  $J$  = 4.8 Hz, 4'-H), 5.10 (s, 1H, 5'-H), 6.62 (dd, 1H,  $J$  = 7.6, 7.0 Hz, 2'-H), 6.82 (d, 4H,  $J$  = 8.9 Hz), 7.20-7.48 (m, 10H, 6-H and aromatic-H), 8.38 (s, 1H, NH).  $^{13}\text{C}$  NMR ( $\text{CDCl}_3$ )  $\delta$  12.69, 37.58 (3'-C), 55.26, 63.15, 67.74, 70.73, 75.24 (4'-

C), 86.01 (2'-C), 86.08, 109.49 (5'-C), 111.80 (5-C), 113.12, 126.74, 127.68, 127.83, 128.24, 130.13, 136.31 (6-C), 145.13, 151.24 (2-C), 158.41, 164.44 (4-C).

A solution of 1.8 g (2.6 mmol) of 11 and  $3 \times 10^{-2}$  g (1.8  $\times 10^{-1}$  mmol) of *N,N*-diisopropylammonium tetrazolide in 50 mL of  $\text{CH}_2\text{Cl}_2$  was treated with 1.1 g (3.6 mmol) of 2-cyanoethoxy-*N,N*-diisopropylphosphoramidite at 23 °C for 16 h. The mixture was then diluted with 100 mL  $\text{CH}_2\text{Cl}_2$  and washed with sat. aqueous  $\text{NaHCO}_3$  (30 mL) and brine (30 mL). The organic layer was dried ( $\text{MgSO}_4$ ) and concentrated at reduced pressure. The residue was purified by  $\text{SiO}_2$  column chromatography (0.5%  $\text{Et}_3\text{N}$  in  $\text{AcOEt}$ ) gave 1.9 g (77.9%) of 12 as a slightly yellow foam.  $^{31}\text{P}$  NMR ( $\text{CDCl}_3$ )  $\delta$  149.33 and 149.53 ppm.

#### 15 EXAMPLE 11

##### 2-(2-Phthalimidoethoxy)ethanol (13)

A suspension of 30.0 g (202.5 mmol) of phthalic anhydride, 20 mL (21.0 g, 199.4 mmol) of 2-(2-aminoethoxy)-ethanol in 400 mL of benzene was added 1.0  $\times 10^{-2}$  g (5.3  $\times 10^{-5}$  mmol) of  $\text{TsOH} \cdot \text{H}_2\text{O}$ . The reaction flask was connected to a Dean-Stark condenser and then heated to reflux for 14 h. The solvent was removed at reduced pressure and the residue was purified by  $\text{SiO}_2$  column chromatography (hexanes/ $\text{AcOEt}$  1:1 then  $\text{CHCl}_3/\text{MeOH}$  20:1) gave 43.7 g (93.1%) of 13 as a light yellow syrup. This syrup slowly crystallized on standing at room temperature overnight as needle-like crystals. mp 66-66.5 °C.  $R_f$  (hexanes/ $\text{AcOEt}$  1:1) 0.17.  $^1\text{H}$  NMR ( $\text{CDCl}_3$ )  $\delta$  2.41 (t, 1H,  $J = 6.2$  Hz, OH), 3.57-3.93 (m, 8 H), 7.69-7.87 (m, 4H).  $^{13}\text{C}$  NMR ( $\text{CDCl}_3$ )  $\delta$  37.52, 61.62, 68.24, 72.24, 123.36, 131.97, 134.00, 168.38 (C=O). Anal Calcd for  $\text{C}_{12}\text{H}_{13}\text{NO}_4$ : C, 61.28; H, 5.53; N, 5.96. Found: C, 61.34; H, 5.62; N, 5.94.

#### EXAMPLE 12A

(2'R,4'S,5'S)-4'-Benzoxo-5'-O-((2-(2-phthalimidoethoxy)-ethyl)-tetrahydrofuran-2'-yl)-*N*<sup>3</sup>-(benzoxymethyl)thymine (14)



A solution of 10.8 g (22.0 mmol) of 8, 10.0 g (42.6 mmol) of 13, and 5.7 g (8.0 mL, 57.0 mmol) of Et<sub>3</sub>N in 200 mL of CH<sub>2</sub>Cl<sub>2</sub> was chilled to -23 °C. To this solution, 16.9 g (14.6 mL, 76.0 mmol) of TMSOTf was added in one portion. The resultant mixture was stirred at -23 °C for 15 minutes and then placed in the freezer (-15 °C) for 16 hours. The reaction was quenched by pouring into 1 L of a 4:1 mixture of AcOEt/H<sub>2</sub>O containing 20 mL of Et<sub>3</sub>N. The organic layer was washed with H<sub>2</sub>O (300 mL), dried (MgSO<sub>4</sub>), and concentrated at reduced pressure. The residue was purified by SiO<sub>2</sub> column chromatography (hexanes/AcOEt 1:1) and gave 12.4 g (84.8 %) of 14 as very hygroscopic white foam. R<sub>f</sub> (hexanes/AcOEt 1:1) 0.22. <sup>1</sup>H NMR (CDCl<sub>3</sub>) δ 1.89 (s, 3H, 5-CH<sub>3</sub>), 2.30 (ddd, 1H, J = 13.2, 8.3, 4.6 Hz, 3'-H<sub>a</sub>), 2.57 (dd, 1H, J = 13.2, 6.5 Hz, 3'-H<sub>b</sub>), 3.63-4.01 (m, 8 H), 4.71 (s, 2H, OCH<sub>2</sub>Ph), 5.19 (s, 1H, 5'-H), 5.36 (d, 1H, J = 4.6 Hz, 4'-H), 5.48 (s, 2H, NCH<sub>2</sub>O), 6.72 (dd, 1H, J = 8.3, 6.5 Hz, 2'-H), 7.28-8.01 (m, 13H, 6-H and aromatic-H), 8.02 (dd, 2H, J = 7.8, 1.5 Hz, aromatic-H). <sup>13</sup>C NMR (CDCl<sub>3</sub>) δ 13.01 (5-CH<sub>3</sub>), 34.87 (3'-C), 36.91, 67.64, 67.78, 69.34, 70.62, 72.05, 73.60 (4'-C), 86.30 (2'-C), 106.16 (5'-C), 111.13 (5-C), 123.08, 127.51, 127.76, 128.21, 128.50, 129.20, 129.70, 131.92, 133.54 (6-C), 134.00, 134.41, 138.08, 151.26 (2-C), 163.07 (4-C), 165.29 (phthalimido C=O) 168.02 (benzoxo C=O). Anal Calcd for C<sub>36</sub>H<sub>35</sub>N<sub>3</sub>O<sub>10</sub> · 0.7 CHCl<sub>3</sub>: C, 58.51; H, 4.74; N, 5.58. Found: C, 58.78; H, 4.74; N, 5.60.

#### EXAMPLE 12B

(2'R,4'S,5'S)-4'-Benzoxo-5'-O-((S)-(2,3-ditetradecanyloxy)-propoxy)-tetrahydrofuran-2'-yl)-N<sup>3</sup>-(benzoxymethyl)thymine (12f)

(S)-1-O-(4-Methoxybenzyl)-2,3-O-isopropylidene-glycerol (12b)

A suspension of (S)-2,2-dimethyl-1,3-dioxolane-4-methanol (5.0 g, 37.8 mmol) and NaH (1.4 g, 58.3 mmol) in DMF (100 mL) was treated with 4-methoxybenzyl chloride (7.7

g, 49.2 mmol) at 0 °C for 4 hours. The resultant mixture was diluted with ether (300 mL) then washed with H<sub>2</sub>O (100 mL). After standard work-up and purification procedure, the title compound was obtained (8.7 g, 91.1%) as a colorless liquid.  
5 bp 152-155 °C/1.1 torr. *R<sub>f</sub>* (hexanes/AcOEt 3:1) 0.44. <sup>1</sup>H NMR (CDCl<sub>3</sub>) δ 1.36 (s, 3H), 1.42 (s, 3H), 3.43 (dd, 1H, J = 9.7, 5.7 Hz, 1-CHH), 3.53 (dd, 1H, J = 9.7, 5.6 Hz, 1-CHH), 3.72 (dd, 1H, J = 8.2, 6.3 Hz, 3-CHH), 3.80 (s, 3H, OCH<sub>3</sub>), 4.05 (dd, 1H, J = 8.2, 6.3 Hz, 3-CHH), 4.29 (m, 1H, 2-H), 4.50  
10 (ABq, 2H, J = 11.7 Hz), 6.86-6.90 (m, 2H), 7.26-7.28 (m, 2H). <sup>13</sup>C NMR (CDCl<sub>3</sub>) δ 25.45, 26.83, 55.20, 66.93, 70.82 (1-C), 73.16 (3-C), 74.80 (2-C), 109.34, 113.82, 129.40, 130.09, 159.32.

**(R)-1-(4-Methoxybenzyloxy)-2,3-dihydroxypropane (12c)**

15 A solution of 12b (8.5 g, 33.7 mmol) in THF (100 mL) was treated with 10% aqueous HCl (10 mL) at 0 °C for 4 hours. The reaction was quenched by adding Et<sub>3</sub>N (20 mL) at 0 °C. After standard work-up and purification procedure, 12c was obtained (7.0 g, 97.9%) as a slightly yellow syrup.  
20 *R<sub>f</sub>* (CHCl<sub>3</sub>/MeOH 10:1) 0.12. <sup>1</sup>H NMR (CDCl<sub>3</sub>) δ 2.45 (br s, 2H, OH), 2.51-3.74 (m, 4H), 3.53 (dd, 1H, J = 9.7, 5.6 Hz, 1-CHH), 3.72 (dd, 1H, J = 8.2, 6.3 Hz, 3-CHH), 3.81 (s, 3H, OCH<sub>3</sub>), 3.85-3.90 (m, 1H), 4.29 (m, 1H, 2-H), 4.87 (s, 2H),  
25 6.86-6.91 (m, 2H), 7.24-7.28 (m, 2H). <sup>13</sup>C NMR (CDCl<sub>3</sub>) δ 55.30, 64.05, 70.87 (2-C), 71.40 (1-C), 73.19 (3-C), 113.89, 129.53, 129.87, 159.36.

**(R)-1-(4-Methoxybenzyloxy)-2,3-(ditetradecanyl-1-oxy)propane (12d)**

30 A suspension of 12c (7.0 g, 33.0 mmol) and NaH (3.2 g, 132.0 mmol) in DMF (100 mL) was treated with 1-bromotetradecane (27.5 g, 99.0 mmol) at 23 °C for 14 hours. The mixture was diluted with AcOEt (500 mL) and washed with H<sub>2</sub>O (100 mL). After standard work-up and purification procedure followed by distillation, 12d was obtained (16.1  
35 g, 80.7%) as a semi-solid wax. bp 244-248 °C/1.5 torr.

$R_f$  (hexanes/AcOEt 40:1) 0.22.  $^1\text{H}$  NMR ( $\text{CDCl}_3$ )  $\delta$  0.88 (t, 6H,  $J$  = 6.7 Hz), 1.26 (m, 48H) 1.51-1.60 (m, 4H), 3.42 (t, 1H,  $J$  = 6.7 Hz, 3-CHH), 3.47-3.56 (m, 8H), 3.81 (s, 3H,  $\text{OCH}_3$ ), 4.48 (s, 2H), 6.85-6.89 (m, 2H), 7.24-7.28 (m, 2H).  $^{13}\text{C}$  NMR ( $\text{CDCl}_3$ )  $\delta$  14.17, 22.76, 26.19, 27.74, 30.17, 32.00, 55.18, 69.99, 70.60, 70.85, 71.67 (1-C), 73.04 (3-C), 77.99 (2-C), 113.72, 129.21, 130.56, 159.17.

**(S)-(2,3-Ditetradecanyloxy)propan-1-ol (12e)**

A solution of 12d (14.0 g, 23.2 mmol) in  $\text{CH}_2\text{Cl}_2$  (180 mL) was added to bromodimethylborane (4.0 g, 32.8 mmol) at -78 °C and the solution was warmed up to 0 °C and stirred for 5 minutes. The resultant mixture was poured into a solution of 4:1 ether/ $\text{H}_2\text{O}$  (500 mL). After standard work-up and purification procedure, 12e was obtained (11.1 g, 98.9%) as a white solid.  $R_f$  (hexanes/AcOEt 9:1) 0.10.  $^1\text{H}$  NMR ( $\text{CDCl}_3$ )  $\delta$  0.88 (t, 6H,  $J$  = 6.7 Hz), 1.26 (m, 48H) 1.51-1.60 (m, 4H), 2.19 (t, 1H,  $J$  = 6.6 Hz, OH), 3.43 (t, 1H,  $J$  = 6.7 Hz, 3-CHH), 3.47-3.78 (m, 8H).  $^{13}\text{C}$  NMR ( $\text{CDCl}_3$ )  $\delta$  14.13, 22.72, 26.13, 29.41, 29.52, 29.70, 30.11, 31.96, 63.00 (1-C), 70.42, 70.88, 71.84 (3-C), 78.41 (2-C).

**(2'R,4'S,5'S)-4'-Benzoxo-5'-O-((S)-(2,3-ditetradecanyloxy)-propoxy)-tetrahydrofuran-2'-yl)-N<sup>3</sup>-(benzoxymethyl)thymine (12f)**

The title compound was prepared according to the procedure of Example 12A. The yield was 93% (light yellow syrup) for the final step.  $R_f$  (hexanes/AcOEt 9:1) 0.26.  $^1\text{H}$  NMR ( $\text{CDCl}_3$ )  $\delta$  0.88 (t, 6H,  $J$  = 6.7 Hz), 1.15-1.34 (m, 48H) 1.51-1.60 (m, 4H), 1.98 (s, 3H, 5- $\text{CH}_3$ ), 2.29 (ddd, 1H,  $J$  = 14.5, 8.3, 4.3 Hz, 3'- $\text{H}_a$ ), 2.62 (dd, 1H,  $J$  = 14.5, 6.2 Hz, 3'- $\text{H}_b$ ), 3.41-3.86 (m, 9H), 4.73 (s, 2H), 5.25 (s, 1H, 5'-H), 5.46 (d, 1H,  $J$  = 4.3 Hz, 4'-H), 5.52 (s, 2H), 6.78 (dd, 1H,  $J$  = 8.3, 6.3 Hz, 2'-H), 7.25-7.65 (m, 9H), 8.02-8.07 (m, 2H).  $^{13}\text{C}$  NMR ( $\text{CDCl}_3$ )  $\delta$  13.21, 14.18, 22.75, 26.17, 29.42, 29.72, 30.10, 31.97, 35.17 (3'-C), 68.72, 70.18, 70.67, 70.93, 71.85, 72.25, 77.52 (4'-C), 86.52 (2'-C), 106.36 (5'-

C), 111.53 (5-C), 127.64, 128.30, 128.57, 129.31, 129.85, 133.66, 134.33 (6-C), 138.10, 151.39 (2-C), 163.30 (4-C), 165.52 (benzoyl C=O). Anal Calcd for  $C_{55}H_{86}N_2O_9$ : C, 71.90; H, 9.37; N, 3.05. Found: C, 71.61; H, 9.24; N, 2.95.

#### 5 EXAMPLE 12C

(2'R,4'S,5'S)-4'-Benzoxymethyl-5'-O-(diethoxyphosphinylmethoxy)-N<sup>3</sup>-(benzoxymethyl)thymine (12g)

The title compound was prepared using diethoxyphosphinylhydroxymethyl (prepared according to the procedure of Phillion, D.P., et.al., Tetrahedron Lett., 1986, 27, 1477) according to the procedure of Example 12A. The yield was 82% (colorless syrup).  $R_f$  ( $CH_2Cl_2$ /AcOEt 3:1) 0.24.  $^1H$  NMR ( $CDCl_3$ )  $\delta$  1.33 (t, 3H,  $J = 7.0$  Hz), 1.39 (t, 3H,  $J = 7.0$  Hz), 2.03 (s, 3H, 5- $CH_3$ ), 2.36 (ddd, 1H,  $J = 14.8, 8.4, 4.9$  Hz, 3'- $H_a$ ), 2.63 (dd, 1H,  $J = 14.8, 6.3$  Hz, 3'- $H_b$ ), 3.84 (dd, H,  $^2J_{gem} = 13.3$ ,  $^2J_{P-H} = 10.0$  Hz, P-CHH-O), 4.02 (dd, H,  $^2J_{gem} = 13.3$ ,  $^2J_{P-H} = 10.0$  Hz, P-CHH-O), 4.18 (qd, 2H,  $^3J_{H-H} = 7.0$  Hz,  $^3J_{P-H} = 0.8$  Hz,  $CH_3CH_2$ -O-P), 4.26 (qd, 2H,  $^3J_{H-H} = 7.0$  Hz,  $^3J_{P-H} = 0.8$  Hz,  $CH_3CH_2$ -O-P), 4.71 (s, 2H), 5.23 (s, 1H, 5'-H), 5.52 (s, 2H), 5.53 (d, 1H,  $J = 4.9$  Hz, 4'-H), 6.86 (dd, 1H,  $J = 8.4, 6.3$  Hz, 2'-H), 7.27-7.66 (m, 9H), 8.02-8.06 (m, 2H).  $^{13}C$  NMR ( $CDCl_3$ )  $\delta$  13.15, 16.72 (d,  $^3J_{P-C} = 5.5$  Hz, O-C-C), 34.67 (3'-C), 61.51 (d,  $^1J_{P-C} = 135.5$  Hz, P-C-O), 63.28 (d,  $^2J_{P-C} = 28.9$  Hz, O-C-C), 70.93, 72.40, 77.41 (4'-C), 86.71 (2'-C), 107.02 (d,  $^3J_{P-C} = 12.8$  Hz, 5'-C), 112.29 (5-C), 127.84, 128.42, 128.85, 129.09, 129.97, 133.91, 138.21, 151.59 (2-C), 163.34 (4-C), 165.63 (benzoyl C=O).  $^{31}P$  NMR ( $CDCl_3$ )  $\delta$  20.45. Anal Calcd for  $C_{29}H_{35}N_2O_{10}P$  0.3  $CHCl_3$ : C, 55.06; H, 5.53; N, 4.39; P, 4.86. Found: C, 54.95; H, 5.58; N, 4.28; P, 4.65.

#### EXAMPLE 12D

(2'R,4'S,5'S)-4'-Benzoxymethyl-5'-O-(benzyloxy)-N<sup>3</sup>-(benzoxymethyl)thymine (12h)

The title compound was prepared using benzyl alcohol according to the procedure of Example 12A. The yield was

90%.

**EXAMPLE 12E**

(2'R,4'S,5'S)-4'-Benzoxo-5'-O-(cyclohexyloxy)-N<sup>3</sup>-(benzoxymethyl)thymine (12i)

5 The title compound was prepared using cyclohexanol according to the procedure of Example 12A. The yield was 93% (white solid). mp 141-3 °C.  $R_f$ (CH<sub>2</sub>Cl<sub>2</sub>/AcOEt 1:1) 0.58. <sup>1</sup>H NMR (CDCl<sub>3</sub>) δ 1.26-1.60 (m, 6H), 1.69-2.04 (m, 4H), 1.96 (s, 3H, 5-CH<sub>3</sub>), 2.31 (ddd, 1H, J = 14.7, 7.9, 4.8 Hz, 3'-H<sub>a</sub>), 2.63 (dd, 1H, J = 14.7, 6.5 Hz, 3'-H<sub>b</sub>), 3.73 (m, 1H), 4.72 (s, 2H), 5.36 (s, 1H, 5'-H), 5.38 (d, 1H, J = 4.8 Hz, 4'-H), 5.52 (s, 2H), 6.75 (dd, 1H, J = 7.9, 6.5 Hz, 2'-H), 7.22-7.65 (m, 9H), 8.02-8.06 (m, 2H). <sup>13</sup>C NMR (CDCl<sub>3</sub>) δ 13.25, 23.74, 23.96, 25.53, 31.83, 33.29, 35.29 (3'-C), 70.69, 72.26, 77.21, 78.26 (4'-C), 86.31 (2'-C), 104.58 (5'-C), 110.88 (5-C), 127.65, 128.31, 128.53, 129.20, 129.84, 133.66, 134.81 (6-C), 138.11, 150.41 (2-C), 163.33 (4-C), 165.74 (benzoyl C=O). Anal Calcd for C<sub>30</sub>H<sub>34</sub>N<sub>2</sub>O<sub>7</sub> · 0.5 H<sub>2</sub>O: C, 66.30; H, 6.45; N, 5.16. Found: C, 66.58; H, 6.44; N, 5.09.

**EXAMPLE 12F**

(2'R,4'S,5'S)-4'-Benzoxo-5'-O-(1-adamantyloxy)-N<sup>3</sup>-(benzoxymethyl)thymine (12j)

25 The title compound was prepared using 1-adamantyl alcohol according to the procedure of Example 12A. The yield was 86% (white solid). mp 98-100 °C.  $R_f$ (hexanes/AcOEt 1:1) 0.60. <sup>1</sup>H NMR (CDCl<sub>3</sub>) δ 1.50-1.95 (m, 12H), 1.98 (s, 3H, 5-CH<sub>3</sub>), 2.05-2.25 (m, 3H), 2.30 (ddd, 1H, J = 14.5, 7.9, 4.5 Hz, 3'-H<sub>a</sub>), 2.63 (dd, 1H, J = 14.5, 6.5 Hz, 3'-H<sub>b</sub>), 4.73 (s, 2H), 5.27 (d, 1H, J = 4.5 Hz, 4'-H), 5.52 (s, 2H), 5.61 (s, 1H, 5'-H), 6.71 (dd, 1H, J = 7.9, 6.5 Hz, 2'-H), 7.20-7.65 (m, 8H), 7.78 (s, 1H, 6-H), 8.08 (d, 2H, J = 7.6 Hz). <sup>13</sup>C NMR (CDCl<sub>3</sub>) δ 13.37, 30.58, 35.42 (3'-C), 42.66, 45.33, 70.65, 72.20, 75.96, 79.14 (4'-C), 86.31 (2'-C), 99.48 (5'-C), 110.54 (5-C), 127.61, 128.28, 129.25,

129.83, 133.61, 135.36 (6-C), 138.14, 151.37 (2-C), 163.36 (4-C), 165.77 (benzoyl C=O). Anal Calcd for  $C_{34}H_{38}N_2O_7$  0.4  $CHCl_3$ : C, 65.13; H, 6.06; N, 4.42. Found: C, 64.97; H, 6.34; N, 4.30.

5 **EXAMPLE 12G**

(2'R,4'S,5'S)-4'-Benzoxo-5'-O-(2-bromoethoxy)-N<sup>3</sup>-(benzoxymethyl)thymine (12k)

The title compound was prepared using 2-bromoethanol according to the procedure of Example 12A. The yield was  
10 91% (white solid). mp 84-6 °C.  $R_f$ (hexanes/AcOEt 3:1) 0.32.  $^1H$  NMR ( $CDCl_3$ )  $\delta$  1.99 (s, 3H, 5- $CH_3$ ), 2.35 (ddd, 1H, J = 14.8, 8.4, 4.9 Hz, 3'- $H_a$ ), 2.63 (dd, 1H, J = 14.8, 6.4 Hz, 3'- $H_b$ ), 3.57 (t, 2H, J = 5.5 Hz,  $BrCH_2CH_2O$ ), 3.99 (dt, 1H, J = 16.4, 5.5 Hz,  $BrCH_2CHHO$ ), 4.10 (dt, 1H, J = 16.4, 5.5 Hz,  $BrCH_2CHHO$ ), 4.72 (s, 2H), 5.25 (s, 1H, 5'-H), 5.49 (d, 1H, J = 4.9 Hz, 4'-H), 5.52 (s, 2H), 6.82 (dd, 1H, J = 8.4, 6.4 Hz, 2'-H), 7.27-7.66 (m, 9H), 8.02-8.06 (m, 2H).  $^{13}C$  NMR ( $CDCl_3$ )  $\delta$  13.42 (5- $CH_3$ ), 30.26 ( $Br-C-C-O$ ), 34.73 (3'-C), 68.87 ( $Br-C-C-O$ ), 70.76, 72.28, 77.48 (4'-C), 86.46 (2'-C),  
15  $BrCH_2CHHO$ ), 106.40 (5'-C), 111.63 (5-C), 127.68, 128.34 (6-C), 128.66, 128.96, 129.86, 133.83, 134.94, 138.05, 151.42 (2-C), 163.24 (4-C), 165.65 (benzoyl C=O). Anal Calcd for  $C_{26}H_{27}N_2O_7Br$ : C, 55.82; H, 4.83; N, 5.01; Br, 14.30. Found: C, 55.78; H, 4.85; N, 4.76, Br, 14.43.  
20

25 **EXAMPLE 12H**

(2'R,4'S,5'S)-4'-Benzoxo-5'-O-(2-propenyl-1-oxy)-N<sup>3</sup>-(benzoxymethyl)thymine (12l)

The title compound was prepared using allyl alcohol according to the procedure of Example 12A. The yield was  
30 84% (white solid). mp 112-3 °C.  $R_f$ (hexanes/AcOEt 1:1) 0.60.  $^1H$  NMR ( $CDCl_3$ )  $\delta$  1.93 (s, 3H, 5- $CH_3$ ), 2.33 (ddd, 1H, J = 14.5, 6.8, 4.7 Hz, 3'- $H_a$ ), 2.64 (dd, 1H, J = 14.5, 8.1 Hz, 3'- $H_b$ ), 4.17 (dd, 1H, J = 12.6, 4.1, 1.0 Hz,  $CHH-CH=CH_2$ ), 4.29 (ddd, 1H, J = 12.6, 5.8, 1.0 Hz,  $CHH-CH=CH_2$ ), 4.72 (s, 2H), 5.24 (s, 1H, 5'-H), 5.30 (ddd, 1H, J = 10.2, 1.3, 1.0  
35

Hz,  $\text{CH}_2\text{-CH=CHH}_{\text{trans}}$ ), 5.30 (ddd, 1H,  $J = 17.3, 1.5, 1.0$  Hz,  $\text{CH}_2\text{-CH=CHH}_{\text{cis}}$ ), 5.47 (d,  $J = 4.7$  Hz, 4'-H), 5.51 (s, 2H), 5.97 (dddd, 1H,  $J = 17.3, 10.2, 5.8, 4.1$  Hz,  $\text{CH}_2\text{-CH=CH}_2$ ), 6.82 (dd, 1H,  $J = 8.1, 6.8$  Hz, 2'-H), 7.18-7.65 (m, 9H),  
 5 8.03-8.06 (m, 2H).  $^{13}\text{C}$  NMR ( $\text{CDCl}_3$ )  $\delta$  13.32 (5- $\text{CH}_3$ ), 35.18 (3'-C), 69.10 (C=C-C), 70.76, 72.29, 77.59 (4'-C), 86.41 (2'-C), 105.52 (5'-C), 111.43 (5-C), 118.29 (C=C-C), 127.62, 128.31 (6-C), 128.60, 129.12, 129.84, 132.94 (C=C-C), 133.70, 134.21, 138.10, 151.43 (2-C), 163.22 (4-C), 165.60  
 10 (benzoyl C=O). Anal Calcd for  $\text{C}_{28}\text{H}_{28}\text{N}_2\text{O}_7$ : C, 68.29; H, 5.69; N, 5.69. Found: C, 68.76; H, 5.63; N, 5.60.

**EXAMPLE 12I**

(2'R,4'S,5'S)-3-Benzoxymethyl-(4'-benzoxy-5'-azidotetrahydrofuran-2'-yl)thymidine

15 The title compound was prepared using TMSN<sub>3</sub> according to the procedure of Example 12A. The yield was 78%.  $R_f$  (hexanes/AcOEt 1:1) 0.36.  $^1\text{H}$  NMR ( $\text{CDCl}_3$ )  $\delta$  1.50-1.95 (m, 12H), 1.98 (s, 3H, 5- $\text{CH}_3$ ), 2.27 (ddd, 1H,  $J = 14.8, 8.7, 4.9$  Hz, 3'-H<sub>a</sub>), 2.66 (dd, 1H,  $J = 14.8, 6.0$  Hz, 3'-H<sub>b</sub>), 4.70 (s, 2H),  
 20 5.37 (d, 1H,  $J = 4.9$  Hz, 4'-H), 5.51 (s, 2H), 5.60 (s, 1H, 5'-H), 6.74 (dd, 1H,  $J = 8.7, 6.0$  Hz, 2'-H), 7.20-7.65 (m, 8H), 7.78 (s, 1H, 6-H), 8.08 (d, 2H,  $J = 7.6$  Hz).  $^{13}\text{C}$  NMR ( $\text{CDCl}_3$ )  $\delta$  13.47, 34.71 (3'-C), 70.80, 72.33, 77.60 (4'-C), 87.14 (2'-C), 95.07 (5'-C), 111.74 (5-C), 127.61;  
 25 127.76, 127.78, 128.31, 128.69, 129.88, 133.29, 133.94 (6-C), 138.04, 151.23 (2-C), 163.04 (4-C), 165.51 (benzoyl C=O). Anal Calcd for  $\text{C}_{24}\text{H}_{23}\text{N}_5\text{O}_6 \cdot 0.45 \text{CHCl}_3$ : C, 55.28; H, 4.42; N, 13.19. Found: C, 55.38; H, 4.39; N, 13.00.

**EXAMPLE 13**

30 (2'R,4'S,5'S)-4'-Hydroxy-5'-O-((2-(2-aminoethoxy)ethyl)-tetrahydrofuran-2'-yl)thymine (15)

A suspension of 12.4 g (18.5 mmol) of 14 and 7.6 g of 10%  $\text{Pd}(\text{OH})_2/\text{C}$  in 100 mL of 1:1 mixture of acetone/MeOH was stirred under an  $\text{H}_2$  atmosphere for 16 hours. The suspended

material was filtered through a pad of Celite and the filtrate was concentrated at reduced pressure to give 7.8 g (76.7%) of de-BOM product as a white form. The de-BOM product was re-dissolved in 100 mL MeOH and 1.54 g (48.0 mmol) of  $\text{H}_2\text{NNH}_2$  was added. The mixture was heated at reflux for 16 hours and then concentrated at reduced pressure. The resultant crude, free amine product (1.8 g) was redissolved in 100 mL MeOH, and 4.0 g (100.0 mmol) of NaOH was then added. The mixture was stirred at 45 °C for 4 hours, and the solvent was then removed at reduced pressure. The residue was purified by  $\text{SiO}_2$  column chromatography (2%  $\text{Et}_3\text{N}$  in  $\text{CHCl}_3/\text{MeOH}$  1:1) to give 2.2 g (37.7% from 14) of 15 as a white solid.  $R_f$  (2%  $\text{Et}_3\text{N}$  in  $\text{CHCl}_3/\text{MeOH}$  1:1) 0.21.  $^1\text{H}$  NMR ( $\text{CD}_3\text{OD}$ )  $\delta$  1.93 (s, 3H, 5- $\text{CH}_3$ ), 2.27 (m, 2H, 3'-H), 2.82 (m, 2H), 3.58-3.93 (m, 6H), 4.30 (d, 1H,  $J = 4.5$  Hz, 4'-H), 5.05 (s, 1H, 5'-H), 6.62 (t, 1H,  $J = 8.1$  Hz, 2'-H), 7.58 (s, 1H, 6-H).  $^{13}\text{C}$  NMR ( $\text{CD}_3\text{OD}$ )  $\delta$  12.80 (5- $\text{CH}_3$ ), 38.33 (3'-C), 51.23, 68.77, 70.00, 71.31, 76.18 (4'-C), 87.29 (2'-C), 110.77 (5'-C), 112.37 (5-C), 137.84 (6-C), 152.75 (2-C), 166.31 (4-C).

#### 20 EXAMPLE 14

(2'R,4'S,5'S)-(4'-Hydroxy-5'-O-((2-aminoethoxy)ethyl)-tetrahydrofuran-2'-yl)thymine, cholesterylforamide (16)

A stirred solution of 2.1 g (6.8 mmol) of 15 in 20 mL of 1:1  $\text{CH}_2\text{Cl}_2/\text{pyridine}$  was added a solution of 3.3 g (7.5 mmol) of cholesteryl chloroformate in 20 mL  $\text{CH}_2\text{Cl}_2$  at 0 °C over 45 min period. The resultant mixture was stirred for additional 3 hours at 0 °C and then diluted with 200 mL  $\text{CH}_2\text{Cl}_2$ , washed with sat. aqueous  $\text{NaHCO}_3$  (50 mL) and brine (50 mL). The organic layer was dried ( $\text{MgSO}_4$ ) and concentrated at reduced pressure. The residue was purified by  $\text{SiO}_2$  column chromatography ( $\text{CH}_2\text{Cl}_2/\text{AcOEt}$  9:1 then  $\text{AcOEt}$ ) to give 2.2 g (45.4%) of 16 as a white solid.  $R_f$  ( $\text{AcOEt}$ ) 0.44.  $^1\text{H}$  NMR ( $\text{DMSO}-d_6$ , assignment for some characteristic proton resonances) :  $\delta$  4.93 (s, 1H, 5'-H), 5.33 (d, 1H,  $J = 3.8$  Hz, 4'-OH), 5.45 (d, 1H,  $J = 3.8$  Hz, 4'-H), 6.48 (t, 1H,  $J = 8.2$  Hz, 2'-H), 8.37 (s, 1H, 6-H), 11.38 (s, 1H, NH).  $^{13}\text{C}$



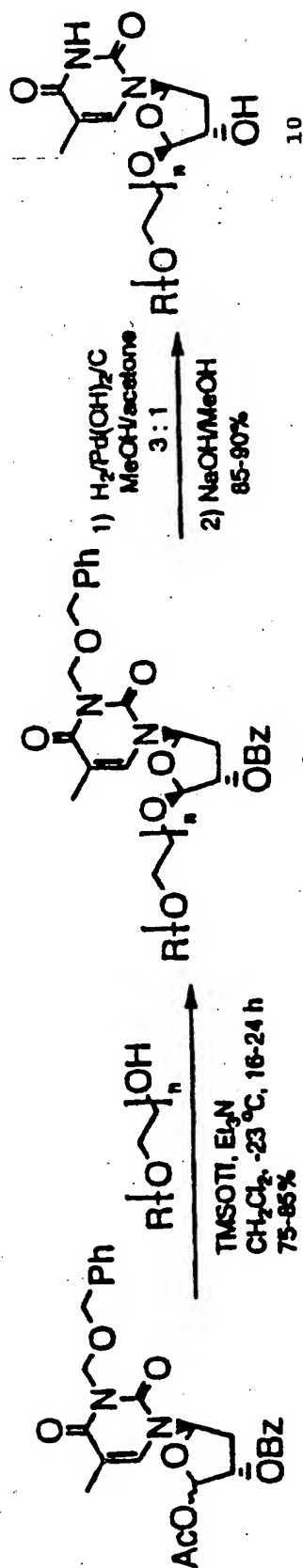
NMR (DMSO- $d_6$ )  $\delta$  11.65, 12.18, 18.55, 18.98, 22.37, 22.61, 23.38, 23.89, 27.43, 27.83, 31.42, 34.72, 34.87, 35.30, 35.78, 36.10, 41.89, (3'-C), 47.38, 49.56, 55.74, 56.20, 66.82, 69.15, 73.01, 73.88, 74.11 (4'-C), 79.20, 85.09 (2'-C), 108.61 (5'-C), 110.34 (5-C), 139.73 (6-C), 150.67 (2-C), 154.99 (formamide C=O), 163.55 (4-C).

#### EXAMPLE 15

(2'R,4'S,5'S)-(4'-O-(2-Cyanoethoxy-N,N-diisopropylamino-phosphorityl)-5'-O-((2-aminoethoxy)ethyl)-tetrahydrofuran-2'-yl) thymine, cholesterylforamide (17)

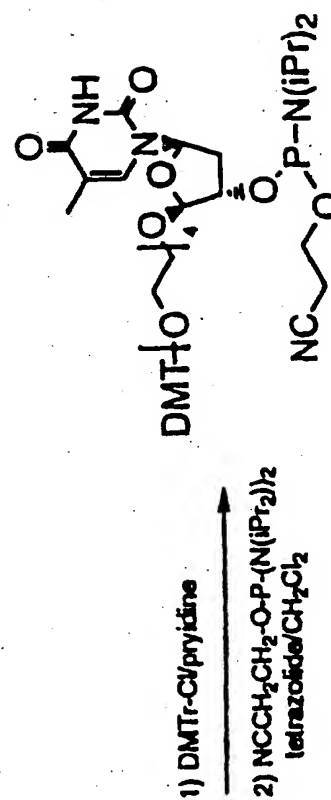
A solution of 1.7 g (2.3 mmol) of 16 and 1.4 g (1.9 mL, 11.0 mmol) of  $i$ -Pr<sub>2</sub>NEt in 40 mL CH<sub>2</sub>Cl<sub>2</sub> was treated with 1.1 g (1.0 mL, 4.6 mmol) of 2-cyanoethyl N,N-diisopropylchlorophosphoramidite at room temperature for 24 hours. The resultant mixture was diluted with 100 mL CH<sub>2</sub>Cl<sub>2</sub> and washed with sat. aqueous NaHCO<sub>3</sub> (30 mL) and brine (30 mL). The organic layer was dried (MgSO<sub>4</sub>) and concentrated at reduced pressure. The residue was purified by SiO<sub>2</sub> column chromatography (0.5% Et<sub>3</sub>N in hexanes/AcOEt 1:1 then AcOEt) to give 0.36 g (16.6%) of 17 as a white foam. <sup>31</sup>P NMR (CDCl<sub>3</sub>)  $\delta$  149.23, 149.43 ppm.

A summary of the synthesis of compounds 8 through 17 are presented on the following pages.



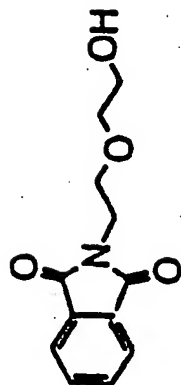
10a: n = 1, R = OH  
10b: n = 3, R = OCH<sub>3</sub>  
10c: n = 4, R = OH  
10d: n = 6, R = OH

9a: n = 1, R = OH  
9b: n = 3, R = OCH<sub>3</sub>  
9c: n = 4, R = OH  
9d: n = 6, R = OH



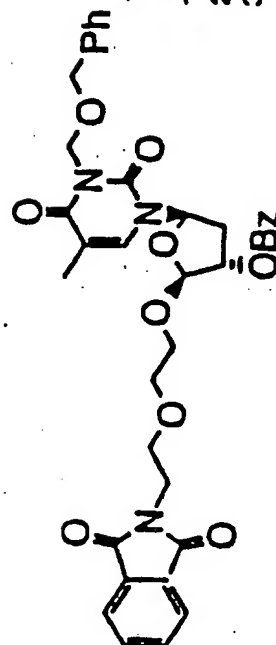


phthalic  
anhydride | p-TsOH, benzene,  
reflux, 82%

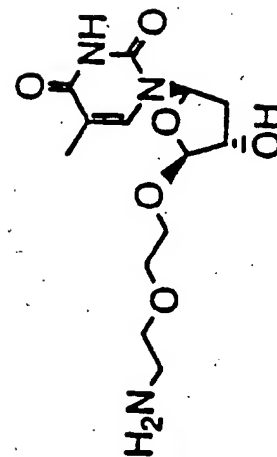


13

TMSOTf, Et<sub>3</sub>N, CH<sub>2</sub>Cl<sub>2</sub>,  
-23 °C, 16-24 h, 93%

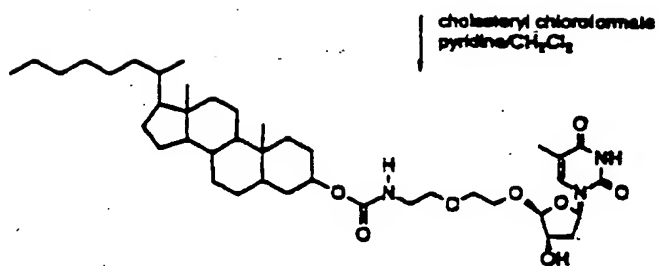
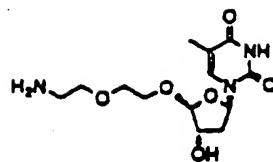


14

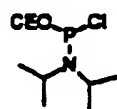


15

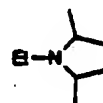
1) H<sub>2</sub>NNH<sub>2</sub>  
MeOH  
2) NaOH/MeOH  
3) H<sub>2</sub>/Pd(OH)<sub>2</sub>/C  
MeOH



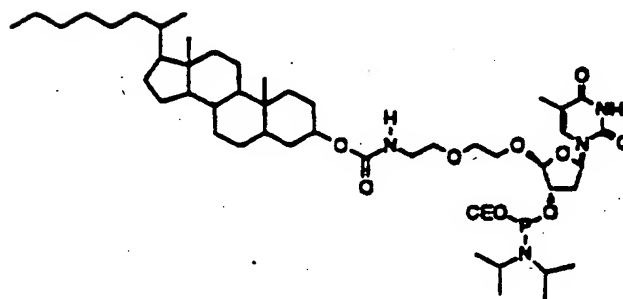
16



(1.4 eq.)



(2.0 eq.)



17

**EXAMPLE 16**

**N<sup>3</sup>-Benzoxyethyl-(4'-O-benzoyl-5'-O-adamantyl-tetrahydrofuran-2'-yl)thymine (18)**

To a solution of 8 (3.5 g, 7.1 mmol) and 1-adamantanol  
5 (2.2 g, 14.2 mmol) in CH<sub>2</sub>Cl<sub>2</sub> (100 ml) was added TMSOTf (2.7  
ml, 3.2 g, 14.2 mmol) at -23 °C in one portion via a  
syringe. The resultant mixture was stirred at -23 °C for  
0.5 hr and then placed in a freezer (-15 °C) for 18 hr. The  
reaction mixture was then poured into a 40:10:1 bilayer  
10 solution of AcOEt/H<sub>2</sub>O/Et<sub>3</sub>N (500 ml), the organic layer dried  
over MgSO<sub>4</sub> and concentrated at reduced pressure. The  
residue was purified by SiO<sub>2</sub> column chromatography  
(hexanes/AcOEt 4:1 then 2:1) to give 3.6 g (86.7%) of the  
title compound as a white solid. R<sub>f</sub> (hexanes/AcOEt 1:1)  
15 0.60. <sup>1</sup>H NMR (CDCl<sub>3</sub>) δ 1.50-1.95 (m, 12 H, adamantyl-H),  
1.98 (s, 3H, 5-CH<sub>3</sub>), 2.05-2.25 (m, 4H, adamantyl-H), 2.3  
(ddd, 1H, J = 14.5, 7.9, 4.5 Hz, 3'-H<sub>a</sub>), 2.63 (dd, 1H, J =  
14.5, 6.5 Hz, 2'-H<sub>b</sub>), 4.73 (s, 2H, OCH<sub>2</sub>Ph), 5.27 (d, 1H, J  
= 4.5 Hz, 4'-H), 5.52 (s, 2H, NCH<sub>2</sub>O), 5.61 (s, 1H, 5'-H),  
20 6.71 (dd, 1H, J = 7.9, 6.5 Hz, 2'-H), 7.2-7.65 (m, 8H,  
aromatic-H), 7.78 (s, 1H, 6-H), 8.08 (d, 2H, J = 7.6 Hz,  
aromatic-H). <sup>13</sup>C-NMR (CDCl<sub>3</sub>) δ 13.37, 30.58, 35.42 (3'-C),  
42.66, 45.33, 70.65, 72.20, 75.96, 79.14 (4'-C) 86.31 (2'-  
C), 99.48 (5'-C), 110.54 (5-C), 127.61, 128.28, 129.25,  
25 129.83, 133.61, 135.36 (6-C), 138.14, 151.37 (2-C), 163.36  
(4-C), 165.77 (benzoyl C=O).

**EXAMPLE 17****Preparation of Oligonucleotides**

30 Oligonucleotides are prepared utilizing normal  
protocol for solid state DNA synthesis, (see *Oligonucleotide  
synthesis, a practice approach*, Ed. M.J. Gait, IRL Press  
1984, Oxford University Press, New York). DMT protected  
phosphoramidite nucleotides are added to the growing  
35 oligonucleotide structure on a solid state support in the  
normal manner. Coupling efficiencies are typically greater  
than 96% for each step of the synthesis with the overall

coupling efficiency greater than 91% for the oligomer. The resulting oligomers are characterized by both gel chromatography and by HPLC using standard protocols.

#### EXAMPLE 18

##### 5 Preparation of Oligonucleotides Containing 5' Terminal Nucleotides Having 4'-Desmethyl Structure Bearing A Conjugation Molecule

An oligonucleotide of the desired sequence is prepared as per the preceding Example 17 in the normal 3' to 5' direction. The last nucleotide of the sequence is preformed to include the desired 4'-desmethyl structure bearing a conjugate group, e.g. compounds 12 and 17, attached via linker groups and compound 18, directly attached, thereon as per the above Examples 10, 15 and 17. This last nucleotide unit, as a 3'-phosphoamidite, is coupled to the remainder of the oligonucleotide sequence utilizing normal protocol for solid state DNA synthesis, as above. As with the remainder of the sequences, coupling efficiencies are generally greater than 96% for this step of the synthesis.

20 A typical synthesis was that of an oligonucleotide of the sequence:

T\*GC ATC CCC CAG GCC ACC AT, SEQ ID NO. 1, where T\* is the nucleotide of compound 12, compound 17 or compound 18 above. The sequences was synthesizer up to and including the penultimate nucleotide unit. The conjugate group bearing, 4-desmethyl, terminal nucleotide, e.g. compound 12, was then used as the ultimate nucleotide unit. The oligonucleotide was removed from the solid support and purified by HPLC in the standard manner.

#### 30 EXAMPLE 18

##### Assay for Separating Expression of Intercellular Adhesion Molecule From Vascular Adhesion Molecules Using Selective Protein Inhibition

Various adhesion molecules are known that are associated with certain adverse events in biological

35

systems. The study of these molecules can be compounded by overlapping activities. Since the intercellular adhesion molecule, ICAM-1, and the vascular adhesion molecule, VCAM-1, both are capable of expressing certain similar properties, e.g. cell adhesion or clumping, segregation of certain of their properties, one from the other, would assist in the evaluation of the properties of a singular adhesion molecule. This cellular assay isolates cytokine induced protein induction of ICAM-1 from that of VCAM-1. The assay differentiates specific protein inhibition of ICAM-1 expression from VCAM-1 expression via use of a probe oligonucleotide having a sequence complementary to the ICAM-1 messenger RNA. Inhibition of the induction of ICAM-1 messenger RNA is effected by treatment of the cells with the probe oligonucleotide in presence of a cationic liposome formulation (DOTMA/DOPE). Cells showing positive adhesion molecule expression in response to cytokine treatment are identified by increases in basal levels of the adhesion molecule protein after treatment with a combination of human TNF- $\alpha$  and murine INF- $\gamma$  using a fluorescence activated cell sorter to identify the cells stained with fluorescent antibodies to either the ICAM-1 or VCAM-1 expressed proteins.

The oligonucleotide utilized as the probe was an oligonucleotide of the invention of the sequence T'GC ATC CCC CAG GCC ACC AT, SEQ ID NO. 1, complementary to the 3' untranslated region of murine ICAM-1 messenger RNA. It includes a conjugation PEG molecule (compound 12 above) as the 5' terminal T nucleotide (nucleotide T' in the above sequence). This oligonucleotide is then used to differentiate ICAM-1 expression from VCAM-1 expression in the test protocol. In cell lines from other species, the base sequence of conjugated oligonucleotide of the invention used in the test is synthesized to be of a sequence complementary to a portion of the ICAM-1 messenger RNA for the species of interest.

A murine endothelioma cell line designated bEND.3,

supplied by Dr. Werner Risau, Max-Planck-Institute, Planegg-Martinsreid, Germany, is tested for adhesion molecule expression as follows. The cells were treated with nanomolar concentrations of the conjugated oligonucleotide, SEQ ID NO. 1, in the presence of 15 µg/ml DOTMA/DOPE liposome (Lipofectin) for 4 hours in a serum free media (Opti-MEM serum-free medium, GIBCO, Grand Island, NY) in the manner described by Chiang et al., *J. Biol. Chem.*, 1991, 266, 18162-18171 or Bennett, et al., *J. Immunol*, 1994, 152, 3530-3540. The media was aspirate and adhesion molecule expression was induced with human TNF-α (5 ng/ml, R&D Systems, Minneapolis, MN) and murine IFN-γ (1000 µ/ml, Genzyme, Cambridge, MA) overnight in DMEM high glucose with 10% fetal bovine serum (Hyclone, Logan, UT). The cells were trypsinized and washed. The cells were stained with ICAM-1-PE fluorescent antibodies (anti-ICAM-1 mAB, PharMingen, San Diego, CA) and VCAM-1-FITC fluorescent antibodies (anti-VCAM-1 mAB, PharMingen, San Diego, CA) and read on a fluorescence activated cell sorter.

Concurrent treatment of the same bEND.3 cell line absent the conjugated oligonucleotide served as a positive control. Both ICAM-1 and VCAM-1 induction were observed in the positive control and normalized to 100%. The conjugated oligonucleotide was used at 0.1, 0.3 and 0.5 µM and the levels of both of induced expression ICAM-1 and VCAM-1 were compared to this positive control. Inhibition of VCAM-1 expression was observed to be 90.08%, 103.82% and 96.33% of control for the 0.1, 0.3 and 0.5 µM concentrations, respectively whereas inhibition of ICAM-1 expression was 3.77%, 4.40% and 8.74% of control for the 0.1, 0.3 and 0.5 µM concentrations, respectively.

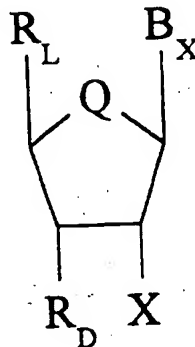
As these results show, in the murine bEND.3 cell line both ICAM-1 and VCAM-1 expression were induced in the positive control. However, in the presence of the ICAM-1 specific messenger RNA complementary oligonucleotide of the invention, VCAM-1 protein expression was maintained but ICAM-1 protein expression was inhibited. Thus the



expression of VCAM-1 is disconnect from that of ICAM-1 such that VCAM-1 expression can be examined independent of ICAM-1 expression.

## WHAT IS CLAIMED IS:

1. A compound of structure:



wherein:

$R_L$  is azido or a group of formula:

5

$R_C - [Y]_e - Z -$

wherein:

Z is O, S or HN;

Y is a bivalent linker;

e is 0 or 1;

10

$R_C$  is alkyl, alkenyl, alkynyl, aryl, alkaryl, aralkyl, phosphinyl, a polyglycol, a polyamine, a polyether, a polythioether, an aromatic or nonaromatic ring system, a steroid molecule, a reporter molecule, an aromatic lipophilic molecule, a non-aromatic lipophilic molecule, a

15

reporter enzyme, a peptide, a protein, a water soluble vitamin, a lipid soluble vitamin, a carbohydrate, a terpene molecule, a phospholipid, an intercalator, a cell receptor binding molecule, a crosslinking agent, or a porphyrin;

$B_X$  is a nucleobase;

20

X is H, OH, O-alkyl, O-alkoxylalkyl, O-alkylamino, or

F;

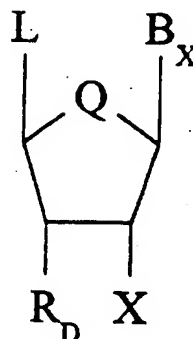
Q is O, S,  $CH_2$ , CHF or  $CF_2$ ; and

$R_D$  is H, hydroxyl, an activated phosphorous group, a nucleoside, an activated nucleotide, a nucleotide, an

25

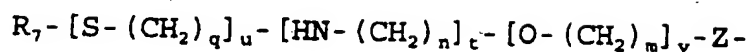
oligonucleotide, an oligonucleoside or a protected derivative thereof.

2. A compound having structure:



wherein:

L is a group of formula:



5 wherein:

t, u and v are each independently integers from 0 to 200;

m, n and q are each independently integers from 1 to 4;

10 Z is S, O or HN;

R<sub>7</sub> is R<sub>C</sub>, H, or a protecting group;

B<sub>x</sub> is a nucleobase;

X is H, OH, O-alkyl, O-alkoxylalkyl, O-alkylamino, or F;

15 Q is O;

R<sub>D</sub> is H, hydroxyl, an activated phosphorous group, a nucleoside, an activated nucleotide, a nucleotide, an oligonucleotide, an oligonucleoside or a protected protected derivative thereof;

20 R<sub>C</sub> is alkyl, alkenyl, alkynyl, aryl, Alkaryl, aralkyl, phosphinyl, a polyglycol, a polyamine, a polyether, a polythioether, an aromatic or nonaromatic ring system, a steroid molecule, a reporter molecule, an aromatic lipophilic molecule, a non-aromatic lipophilic molecule, a  
25 reporter enzyme, a peptide, a protein, a water soluble vitamin, a lipid soluble vitamin, a carbohydrate, a terpene

molecule, a phospholipid, an intercalator, a cell receptor binding molecule, a crosslinking agent, or a porphyrin; and provided that t, u and v are not all simultaneously 0.

- 5           3.     A compound of claim 2 wherein t and u are 0.
4.     A compound of claim 2 wherein t and u are 0, and  
          m is 2.
5.     A comound of claim 2 wherein u and v are 0.
6.     A compound of claim 2 wherein u and v are 0 and  
10       n is 2.
7.     A compound of claim 2 wherein X is H or OH.
8.     A compound of claim 2 wherein R<sub>b</sub> is H or OH.
9.     A compound of claim 2 wherein R<sub>b</sub> is an activated  
          phosphorous group.
- 15          10.    A compound of claim 2 wherein R<sub>b</sub> is an  
          oligonucleotide.
11.    A comound of claim 2 wherein t and v are 0.
12.    A compound of claim 2 wherein t and v are 0 and  
          q is 2.
- 20          13.    A compound of claim 2 wherein Z is O or N.
14.    A compound of claim 2 wherein R<sub>c</sub> is a steroid  
          molecule.
15.    A compound of claim 14 wherein the steroid  
          molecule is cholic acid, deoxycholic acid, dehydrocholic

acid, cortisone, digoxigenin, testosterone, cholesterol or 3-trimethylaminomethylhydrazido cortisone.

16. A compound of claim 2 wherein  $R_c$  is a water soluble vitamin.

5        17. A compound of claim 16 wherein the water soluble vitamin is thiamine, riboflavin, nicotinic acid, pyridoxal phosphate, pyridoxine, pyridoxamine, deoxypyridoxine, pantothenic acid, biotin, folic acid, 5'-deoxyadenosylcobalamin, inositol, choline or ascorbic acid.

10        18. A compound of claim 2 wherein  $R_c$  is a lipid soluble vitamin.

19. A compound of claim 18 wherein the lipid soluble vitamin is a retinal, a retinol, retinoic acid,  $\beta$ -carotene, vitamin D, cholecalciferol, a tocopherol, or a phytol.

15        20. A compound of claim 2 wherein  $R_c$  is a protein.

21. A compound of claim 20 wherein the protein is a phosphodiesterase, a peroxidase, a phosphatase or a nuclease.

20        22. A compound of claim 2 wherein  $R_c$  is a reporter molecule.

23. A compound of claim 22 wherein the reporter molecule is a chromophore, a fluorophore or a radiolabel-containing moiety.

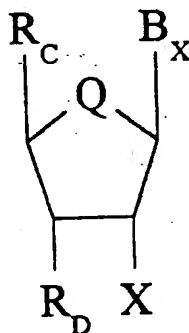
25        24. A compound of claim 23 wherein the fluorophore is fluorescein, chrysine, anthracene, perylene, pyrene, rhodamine.

25. A compound of claim 2 wherein X is F, O-alkyl

having from one to six carbons, O-alkoxyalkyl having from 2 to 6 carbons, or O-alkylamino having from one to six carbons.

26. A compound of claim 2 wherein  $R_D$  is H or OH.

5 27. A compound having structure:



wherein:

$R_C$  is O-alkyl, O-alkenyl, O-alkynyl, O-aryl, O-alkaryl, O-aralkyl, a polyglycol, an aromatic or nonaromatic ring system, a polyamine, a polyether, a steroid molecule, a reporter molecule, an aromatic lipophilic molecule, a non-aromatic lipophilic molecule, a reporter enzyme, a peptide, a protein, a water soluble vitamin, a lipid soluble vitamin, a carbohydrate, a terpene molecule, a phospholipid, an intercalator, a cell receptor binding molecule, a crosslinking agent, or a porphyrin;

$B_X$  is a nucleobase;

$X$  is H, OH, O-alkyl, O-alkoxylalkyl, O-alkylamino, or F;

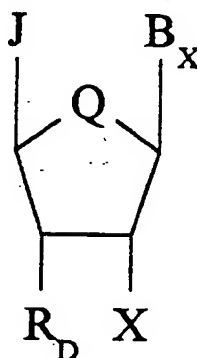
$Q$  is O, S,  $CH_2$ , CHF or  $CF_2$ ; and

20  $R_D$  is H, hydroxyl, an activated phosphorous group, a nucleoside, an activated nucleotide, a nucleotide, an oligonucleotide, an oligonucleoside or a protected protected derivative thereof.

28. A compound of claim 27 wherein  $R_C$  is a

polyether, a polyamine or a non-aromatic lipophilic molecule.

29. A compound of structure:



wherein:

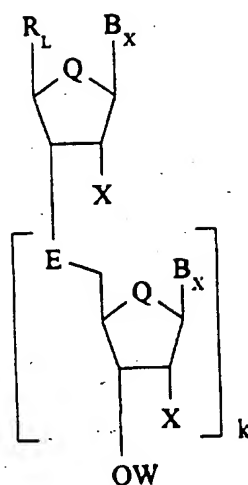
- 5        J is a leaving group;  
       B<sub>x</sub> is a nucleobase;  
       X is H, OH, O-alkyl, O-alkoxylalkyl, O-alkylamino, or  
 10        F;

      Q is O, S, CH<sub>2</sub>, CHF or CF<sub>2</sub>; and

- 10        R<sub>D</sub> is H, hydroxyl, an activated phosphorous group, a  
       nucleoside, an activated nucleotide, a nucleotide, an  
       oligonucleotide, an oligonucleoside or a protected  
       derivative thereof.

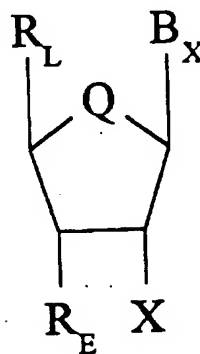
30. A compound of claim 29 wherein J is OH, SH, NH<sub>2</sub>,  
 15        trifluoromethylsulfonyl, methylsulfonyl, halogen, O-  
       trichloroacetimidate, acyloxy, dialkyl phosphite, 2,4,6-  
       trichlorophenyl, p-toluenesulfonyl, 4-dimethylaminoazo-  
       benzenesulfonyl or 5-dimethylaminonaphthalenesulfonyl.

31. A method for forming a 5'-desmethyl conjugate  
 20        oligomer having structure:



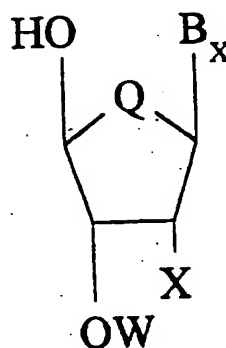
comprising the steps of:

(a) providing a first synthon having structure:



and

(b) contacting the first synthon with a second  
5 synthon having structure:

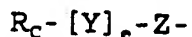


the contacting being for a time and under reaction  
conditions sufficient to form the conjugated oligomer;



wherein:

$R_L$  is azido or a group of formula:



Z is O, S or HN;

5 Y is a bivalent linker;

e is 0 or 1;

$R_C$  is alkyl, alkenyl, alkynyl, aryl, alkaryl, aralkyl, phosphinyl, a polyglycol, a polyamine, a polyether, a polythioether, an aromatic or nonaromatic ring system, a steroid molecule, a reporter molecule, an aromatic lipophilic molecule, a non-aromatic lipophilic molecule, a reporter enzyme, a peptide, a protein, a water soluble vitamin, a lipid soluble vitamin, a carbohydrate, a terpene molecule, a phospholipid, an intercalator, a cell receptor binding molecule, a crosslinking agent, or a porphyrin;

15 E is a phosphodiester, phosphorothioate, phosphorodithioate, phosphamidate, alkyl phosphonate or phosphotriester phosphorous linking group;

$R_E$  is an activated phosphorous group;

20 W is H, a hydroxyl protecting group or a solid state support;

k is an integer from 0 to 50;

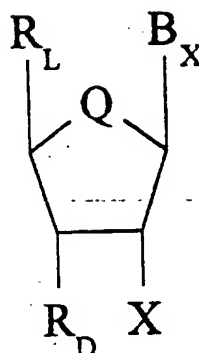
$B_x$  is a nucleobase;

25 X is H, OH, O-alkyl, O-alkoxylalkyl, O-alkylamino, or F;

Q is O, S,  $CH_2$ , CHF or  $CF_2$ ; and

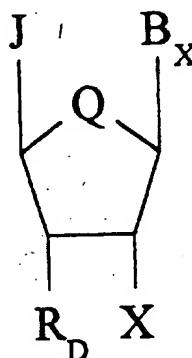
$R_D$  is H, hydroxyl, an activated phosphorous group, a nucleoside, an activated nucleotide, a nucleotide, an oligonucleotide, an oligonucleoside or a protected derivative thereof.

30 32. A method for forming a 5'-desmethyl conjugated monomer having structure:



comprising the steps of:

(a) providing a synthon having structure:



and

(b) contacting this synthon with an activated  
5 conjugating group;

the contacting being for a time and under reaction  
conditions sufficient to form the conjugated monomer;

wherein:

$R_L$  is a group of formula:

10

$R_C - [Y]_e - Z -$

$Z$  is O, S or HN;

$Y$  is a bivalent linker;

$e$  is 0 or 1;

15  $R_C$  is alkyl, alkenyl, alkynyl, aryl, alkaryl, aralkyl,  
phosphinyl, a polyglycol, a polyamine, a polyether, a  
polythioether, an aromatic or nonaromatic ring system, a  
steroid molecule, a reporter molecule, an aromatic  
lipophilic molecule, a non-aromatic lipophilic molecule, a

reporter enzyme, a peptide, a protein, a water soluble vitamin, a lipid soluble vitamin, a carbohydrate, a terpene molecule, a phospholipid, an intercalator, a cell receptor binding molecule, a crosslinking agent, or a porphyrin;

5 J is a leaving group;

B<sub>x</sub> is a nucleobase;

X is H, OH, O-alkyl, O-alkoxylalkyl, O-alkylamino, or F;

Q is O, S, CH<sub>2</sub>, CHF or CF<sub>2</sub>; and

10 R<sub>p</sub> is H, hydroxyl, an activated phosphorous group, a nucleoside, an activated nucleotide, a nucleotide, an oligonucleotide, an oligonucleoside or a protected derivative thereof.

33. The method of claim 32 wherein J is OH, SH, NH<sub>2</sub>,  
15 trifluoromethylsulfonyl, methylsulfonyl, halogen, O-trichloroacetimidate, acyloxy, dialkyl phosphite, 2,4,6-trichlorophenyl, p-toluenesulfonyl, 4-dimethylaminoazobenzenesulfonyl or 5-dimethylaminonaphthalenesulfonyl.

34. The method of claim 32 wherein J is  
20 trifluoromethylsulfonyl, methylsulfonyl, halogen or acyloxy.

## INTERNATIONAL SEARCH REPORT

International application No.  
PCT/US95/13038

## A. CLASSIFICATION OF SUBJECT MATTER

IPC(6) : C07H 21/00, 21/02, 21/04

US CL : 536/24.1

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 536/24.1, 23.1, 24.3, 24.5, 25.3, 25.32; 514, 44; 435, 6

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

Please See Extra Sheet.

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	US, A, 5,223,618 (COOK ET AL) 29 June 1993, see entire document.	1-34
Y	US, A, 4,908,453 (COCUZZA) 13 March 1990, see entire document.	1, 2, 16, 17, 27, 29, 31, 32
Y	US, A, 4,958,013 (LETSINGER) 18 September 1990, see entire document.	1, 2, 14, 15, 27, 29, 31, 32
Y	US, A, 5,262,536 (HOBBS, JR.) 16 November 1993, see entire document.	1, 2, 22-24, 27, 29, 31, 32
P,Y	US, A, 5,391,723 (PRIEST) 21 February 1995, see entire document.	1, 2, 22-24, 27, 29, 31, 32

☒ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:	T	later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
*A* document defining the general state of the art which is not considered to be of particular relevance	X	document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
*E* earlier document published on or after the international filing date	Y	document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
*L* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	A	document member of the same patent family
*O* document referring to an oral disclosure, use, exhibition or other means		
*P* document published prior to the international filing date but later than the priority date claimed		

Date of the actual completion of the international search

11 JANUARY 1996

Date of mailing of the international search report

01 FEB 1996

Name and mailing address of the ISA/US  
Commissioner of Patents and Trademarks  
Box PCT  
Washington, D.C. 20231

Facsimile No. (703) 305-3230

Authorized officer

THOMAS G. LARSON, PH.D.

Telephone No. (703) 308-0196

## C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	Bioconjugate Chemistry, Vol. 1, No. 3, issued May/June 1990, Goodchild, "Conjugates of Oligonucleotides and Modified Oligonucleotides: A Review of their Synthesis and Properties," pages 165-186, see entire document, but especially pages 170 and 173-174, and Table VI.	1-34
Y	Chemical Reviews, Vol. 90, No. 4, issued 1990, Uhlmann et al, "Antisense Oligonucleotides: A New Therapeutic Principle," pages 544-584, see entire document, but especially pages 558-559 and Figure 35C.	1-34
Y	Bioconjugate Chemistry, Vol. 2, No. 6, issued 1991, Tung et al, "Preparation of Oligonucleotide-Peptide Conjugates," pages 464-465, see entire document.	1-34

**B. FIELDS SEARCHED**

Electronic data bases consulted (Name of data base and where practicable terms used):

Data Bases Searched: APS, STN (BIOSIS, CAPLUS, CJACS, DISSABS, MEDLINE, WPIDS)

Search Terms: oligonucleotide(s), ribooligonucleotide(s), deoxyribooligonucleotide(s), nucleoside(s) deamethyl, des methyl, furanyl, steroid, sterol, cholesterol, vitamin(s), folic acid, biotin, conjugate, linker, reporter, label; Cook, Phillip, Dan; Cook, Dan, P; Cook, P. Dan; Cook, P D; Tseng, Kelly; Tseng, K.

Structure Search: Compounds 14, 15, 16

**This Page is Inserted by IFW Indexing and Scanning  
Operations and is not part of the Official Record**

## **BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ BLACK BORDERS
- ☐ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- ☐ FADED TEXT OR DRAWING
- ☐ BLURRED OR ILLEGIBLE TEXT OR DRAWING
- ☐ SKEWED/SLANTED IMAGES
- ☐ COLOR OR BLACK AND WHITE PHOTOGRAPHS
- ☐ GRAY SCALE DOCUMENTS
- ☒ LINES OR MARKS ON ORIGINAL DOCUMENT
- ☐ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- ☐ OTHER: \_\_\_\_\_

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.**

**THIS PAGE BLANK (USPTO)**